The Dissertation Committee for Alvaro Ariel Quezada-Hofflinger Certifies that this is the approved version of the following dissertation:

DO COMPETITION AND ACCOUNTABILITY IMPROVE QUALITY OF EDUCATION? THE CHILEAN CASE FROM 2002 TO 2013

Committee:

James K. Galbraith, Supervisor

Paul von Hippel, Co-Supervisor

Peter Ward

Chandler Stolp

Robert Gilbert

Do Competition and Accountability Improve Quality of Education? The Chilean case from 2002 to 2013

by

Alvaro Ariel Quezada-Hofflinger, BA;MA

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas at Austin

August, 2015

Dedication

To Elisa Hofflinger

Acknowledgements

It has been a long trip since I left my nice and comfortable home in *Huillinlebu*, a small rural community in the south of Chile. At the time, we were 12 students from 1st to 6th grade in one tiny school called *Escuela G-230 Huillinlebu*, which basically was one classroom with two blackboards for all the students. We just had one teacher. We did not take any standardized tests. Instead, we were focused on learning and exploring nature. I would like to thank all these farmers from *Huillinlebu*, this work has been highly inspired by them. After my school day I used to walk to my grandfather's or my next-door neighbor's house. I would like to specifically acknowledge them: my grandfather José, uncle Reinaldo, Elizabeth and their three sons, Braulio, Andres and Alexis. My next-door neighbors who took care of me when my mom passed away, Don Domingo Salinas and Doña Francisca Ramirez, and their three children: Marcela, Enrique and Johanna.

I could not start or finish this PhD without the support of my advisor James K. Galbraith, who helped me since I took his class "Inequality" during my master's program, and which connected all the dots. Furthermore, I would like to thank Paul von Hippel, my co-advisor, who gave me financial support for more than three years and more importantly helped me with details and the big picture of this dissertation. He is still giving good comments and feedback about how I can improve my work. Also I am very grateful to Chandler Stolp, who has been my mentor in quantitative methods and has given me his advice when I got lost, always with a compass and map to show me the right path. Last but not least, I would like to mention Peter Ward who encouraged me to spend a semester in Chile and add qualitative methods to my work, which really changed and improved this dissertation. Furthermore, Robert Gilbert who allowed me to take his class "decision analysis" and be part of my committee. Finally, I would like to

acknowledge the help of Chilean government and especially the Ministry of Education for allowing me to use its datasets, without them I could not finish this project as well as I would like recognize all these teachers, policy makers and academics who participated in this study. Thank you for the time and support.

I would like to acknowledge all the friends that I have met here at UT, those who already finished their programs and left Austin: Jaime Alves, Sandra Botero, Nestor Castañeda, Laura Gamboa, Teresa Correa, Sebastian Valenzuela, Javier Morales, Denisse Gelber, Laura Spanoglo and among many others. And those who are still around and belong to the Zacapa World: Jorge Derpic (Diavolo), Maria Jose Afanador (La reina), Andres Lombana (El psiquico), Edwin Roman (Johnny), Eva Hershaw (Evitala), Juan Sequeda (Wikileaks) and Luciano Correa (Matador). Especially thanks to Marcelo Somos (Maximus) who has been great friend in the good and bad times in Austin and Luis Rios (Lucho) who always has been there for me since we met in the cold winter in Patagonia (Coyhaique), always with a smile and warm hug.

I am grateful to my family, my father Desiderio, my mom Elisa and brothers Matias and Fabian. I would like especially like to thank my wonderful wife, Lindsey, for all her love and support all these years. Finally, I would like to thank all the farmers, all those students who are still walking to their schools in the rain during winter time, all these simple and nice people from my second rural home of "Selva Oscura," who have made this trip with me and help me to keep focused on what really matters in life. Finally, I express thanks to all my classmates from the public boarding schools where I attended middle school (*San Pedro F-218*) and high school (*Liceo B-10),* with whom I shared a roof and a dream of a better future for our community.

Do Competition and Accountability Improve Quality of Education? The Chilean case from 2002 to 2013

Alvaro Ariel Quezada-Hofflinger, Ph.D.

The University of Texas at Austin, 2015

Supervisor: James K. Galbraith

Co-Supervisor: Paul von Hippel

This dissertation studies the macro-level impact of market-oriented reforms in education. Specifically, it evaluates the intended and unintended consequences of the introduction of competition and the high-stakes accountability program in Chile. Competition and high-stakes accountability systems in education are neoliberal, market-oriented policies implemented by lawmakers with the stated goal of improving the quality of education. Performance on standardized tests is a key marker of quality in this system—schools are ranked based on student scores. Funding opportunities are also attached to performance. Thus, advocates argue that if schools compete with each other and parents have the freedom and information to choose their children's schools, the education system would react to these pressures and lead schools with lower performance on standardized tests to eventually close due to low enrollment. Therefore, the overall quality of education would improve.

Influenced by economist Milton Friedman and under the military regimen of August Pinochet, Chile implemented a universal education voucher program in 1981. In 2015, more than half of Chilean students in primary and secondary education are enrolled in private-voucher schools. Furthermore, and as part of its neoliberal agenda, the Chilean government also implemented a high-stakes accountability program as way to make teachers and school administrators responsible for student performance on standardized test.

I use a mixed methods research approach to explore the intended and unintended consequences of the introduction of competition and the high-stakes accountability system. The results show competition has had a negative impact on quality of education at the national level, while families and students from Santiago, the capital, have benefited from competition. Furthermore, in those schools that participate in the high-stakes accountability system, contrary to the expected outcome, teachers are not increasing their use of academic strategies, such as spending more time with students, finding new learning methods, or giving students more homework or assignments in order to improve their performance on standardized tests. Instead, teachers and schools are increasingly using non-academic strategies, such as excluding low-performing students from the test-taking pool, as a way to improve scores.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Around the world, education is overwhelmingly conducted in the public sector and one of the central debates among governments is how to improve the quality of public schools (Kremer & Sarychev, 2000). Politicians and policymakers have proposed incorporating the private sector and market rules into the educational arena as a way to improve the quality of public education. Market-oriented reform in education combines competition, free choice, and accountability as complements to one another. It is assumed that if schools compete with each other and if parents have the freedom to choose their children's school, the educational system will react to these pressures and quality will improve. Parents would be inclined to choose schools with the best quality of education for their children, which would lead those schools with lower quality to eventually close due to low enrollment. Moreover, proponents of school choice argue that in the traditional educational system, principals and teachers do not face strong enough consequences for the poor performance of their schools or students (Betebenner, Howe, & Foster, 2005). They believe the lack of competition and information about the quality of the education a school provides means that schools have few incentives to boost their performance. Instead, an educational system that incorporates competition among schools and in which teachers who are responsible for students' performance on standardized tests is viewed as more likely to improve social mobility and equalize the quality of education among families of different socio-economic backgrounds (Haertel & Herman, 2005; Hamilton, Stecher, & Klein, 2002; Hursh, 2007; Wall, 2000; West, 1997).

However, while the privatization of education is widely advocated from a neoliberal standpoint, hardly any empirical evidence exists in support of this position, at least not at the macro level. In fact, discussions on such reforms are mostly based on

ideological beliefs (D. E. Campbell, West, & Peterson, 2005; Levin, 1998). To fill this lacuna in the research, this study assesses the outcomes of market-oriented reform in education, especially the macro-level impact of the introduction of competition and accountability systems. It does so by analyzing the case of Chile, where a large-scale, universal educational voucher system was implemented during the 1980's and a high-stakes accountability program in 2008.

I apply a mixed-methods research approach that draws heavily on a quantitative analysis of nation-wide education and standardized test data, and incorporates semi-structured interviews to meet three objectives. The first objective explores the relationship between school competition and student achievement in the period between 2002 and 2013. During this period, the Chilean educational system experienced an important increase in students enrolled in private-voucher schools—the percentage of students in primary and secondary education enrolled in private-voucher schools increased from 36.4 percent in 2000 to 55.5 percent in 2013. To meet this objective, I evaluate the neoliberal hypothesis that in a competitive educational system parents will choose the best school for their children based on objective information that standardized tests provide on school quality. Schools with poorer standardized test scores would lose students and eventually disappear from the market and the whole system would improve as a result of that, ultimately improving the quality of the education in general. Furthermore, in a competitive educational system, the quality of schools should improve regardless of their location and under what conditions they exist because the whole system would be improved by market-based reform. Thus, my exploration of this hypothesis is guided by the following sub-questions: (1.1) how does competition affect students' achievement? (1.2) does competition have similar impacts on the performance of students who are educated in the capital city, Santiago, and outside of it?

A second objective of this research is to understand the intended and unintended consequences of the introduction of the accountability system in Chile, and in particular what are teachers' responses to the implementation of the high-stakes accountability system and what incentives exist to game the system or improve students' leaning. A second wave of market-oriented policies in education in Chile, after the introduction of competition through private-voucher schools in 1981, was the implementation of a high-stakes accountability system in 2008. Previous studies have shown that those schools that participate in the high-stakes accountability program have increased student test scores (Correa, Parro, Reyes, & Ugarte, 2012; Mineduc, 2012; Mizala & Torche, 2013; Perticara, Román, & Selman, 2013; Villarroel, 2012). However, these studies do not reveal what schools did in order to improve their performance. Schools can increase students' scores in several positive or negative ways, including "gaming the system" or by improving student learning, both strategies better scores. In evaluating this hypothesis, this study asks: (2.1) have teachers increased the use of academic strategies as a way to improve students' performance on standardized tests after the implementation of the high-stakes accountability system in 2008?; (2.2) what are the difficulties and constraints that the Chilean educational system imposes on teachers and school administrators in order to increase the use of academic strategies under the high-stakes accountability system?

The third objective of this dissertation analyzes the unintended responses of teachers under the implementation of high-stakes accountability program. In particular, I study the aftermath of the implementation of the high-stakes accountability system in 2008 and whether or not low-achieving students are more likely to be excluded from the test taking pool. Another goal is to establish whether or not the new accountability system had a meaningful impact on the schools' ranking due to the exclusion of low-

achieving students. I draw upon Campbell's law, which points out that "the more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social process it is intended to monitor" (Campbell, 1979). I use Campbell's law to study the impact of using a standardized test, as a single indicator, to measure the quality of education before and after the implementation of the high-stakes accountability system in Chile. In particular, I ask the following questions: (3.1) are schools that participate in the high-stakes accountability program more likely to exclude low-performing students from the test-taking pool in order to improve their overall scores on standardized tests? (3.2) how do the schools' rankings based on standardized tests change if the students excluded from the test-taking pool were able to take the standardized test?

SIGNIFICANCE

The privatization of educational systems remains an understudied topic largely due to measurement problems and a lack of data. While school choice programs are fairly common, they have been implemented in unique contexts with very different geographic and socioeconomic characteristics. Moreover, in the cases where data are available, scholars have focused on comparing the performance difference between public and voucher schools at the school or student level, and have not devoted as much attention to the effects of voucher school programs at the macro level (Torche, 2005). The Chilean case provides a rich opportunity for research because its voucher system has operated nationwide, almost unaltered, for more than thirty years, and Chile's extensive database offers an ideal platform for analyzing the impact of private-voucher school systems (Carnoy & McEwan, 2000). In addition, the Chilean case is interesting because many other Latin American countries implemented neoliberal policies, such as voucher school

programs, during the 1990s. As such, the Chilean experience can offer valuable lessons to countries in the region and the topic is a high priority for the Chilean government since the student movement (2011 & 2012) demanded regulation of the market-driven Chilean educational system and the strengthening of public education.

There is virtually no evidence at the aggregate level of the effects of privatization reform resulting from large-scale voucher programs (Nechyba, 2000). As such, given that the question is still relevant, U.S. researchers and policymakers should consider the evidence from large-scale programs in other countries such as Chile and New Zealand more carefully (Ladd, 2002). In this sense, there are lessons to be learned from the Chilean experience, not only for countries in Latin America, but also for policymakers in the U.S. As a result, policymakers, scholars and practitioners all stand to benefit from an in-depth, systematic assessment of its actual impact on educational outcomes.

This study also makes a contribution to the current research on Chile in terms of data and the methodology used. First, the data analyzed is qualitatively better than the data utilized in previous research because the period analyzed is longer and incorporates a richer level of information on students, students' family characteristics, teachers, schools and school districts. Moreover, the data includes all 345 municipalities, whereas in most studies, small school districts are excluded or underrepresented in the national studies, even though they educate the most disadvantaged populations. Second, to estimate the relationship between competition and student achievement rather than use instrumental variables to deal with endogeneity problems with competition and test scores, I used a fixed effects model at municipality level. In this way a municipality is compared with itself through time and it is possible to obtain unbiased estimation of the effect of competition on students' achievement.

Finally, in order to study how teachers respond after the implementation of a high-stakes accountability program previous studies in Chile have used qualitative methodologies (Falabella, 2013), but do not examine what type of municipalities, schools, labor conditions, or leadership help to develop academic or non-academic strategies to improve students' scores on standardized tests. In this sense, this dissertation offers fundamental information to policy-makers in Chile about the intended and unintended consequences of market-oriented policies implemented in Chile, in particular the impact of competition and high-stakes accountability system on student achievement. This information is crucial because until 2014 the high-stakes program was voluntary, but most likely in 2015 the program will become mandatory for all schools and will be implemented nationwide. Thus, the results of the dissertation will help inform the Ministry of Education's efforts to either reduce or ameliorate the unintended consequences of this type of program and reinforce the intended effect of this public policy.

BACKGROUND

This study focuses on the Chilean case due to that country's long-standing implementation of market-based reforms to its education system. Under the military regime of Augusto Pinochet (1973-1990), the Chilean educational system underwent one of the most radical changes of its history as part of a market-orientated transformation undertaken during the 1980's (Torche, 2005). Influenced by the economist Milton Friedman's ideas, Chilean policy-makers advocated for a reduction in the government's role in education arguing that free market competitors replace the government monopoly, Chile became one of the first countries in the world to implement such reforms. Even though the democratic government, which took office in 1990, disagreed with a number

of the military regime's educational policies, it did not introduce substantive changes to the private-voucher model. The Chilean educational system continues to be almost the same as the one that Pinochet implemented in 1981. The reform of the educational system that took place during the Pinochet dictatorship had two main components: (1) the decentralization of education and (2) the creation of a universal educational voucher system.

Decentralization of Education

Traditionally in Chile, the Ministry of Education centrally controlled all public schools and was responsible for all aspects of their operation, such as teacher pay, facility maintenance, and curriculum design. Pinochet's regime transferred all public schools, which were under the administration of the Ministry of Education, to the municipalities. With the reform, school control was transferred to municipalities and the reform strongly decreased the national government's role and expenditures with respect to education. For example, state funding for education was reduced from 7.2 percent of the GDP in 1972 to 2.4 percent in 1990 (Claro, 2005).

Private-Voucher Schools

Another key component of the privatization reform was the introduction of a universal educational voucher system. Within this system, a subsidy is paid to public and private non-voucher schools on the basis of student enrollment, and parents/legal guardians are free to choose which school they would like their children to attend (Torche, 2005). For example, if a student is absent three days during the month, this money is discounted from the government's monthly payment to the school. According to the free market reforms implemented during the eighties in Chile, private schools can decide to increase their tuition according to their own interests and priorities. Therefore,

vouchers only cover a certain percentage of tuition, and parents need to pay extra money to register their children in these schools. Moreover, the schools can select their students through entrance exams, according to their own criteria. The government does not intervene to control the tuition rates of private-voucher schools or in the process of selection.

In this sense, it is important to note that the Chilean voucher system differs significantly from the Wisconsin, New York or Florida system: the Chilean government does not give a tuition certificate to the family but, rather, pays the subsidy directly to the school that the student or parent chooses. Thus, the Chilean voucher model is known as a "funds follow the student" system (Mizala & Romaguera, 2000), where the government pays each school a specific amount of resources for each student that successfully attends classes. Each school's revenue is determined on a month-to-month basis as a function of the schools' total enrollment (Carnoy & McEwan, 2000). Some of the schools that emerged after the privatization reform were managed by religious and non-profit organizations, but the majority of them were run by private agents that capitalized on education as a profitable business (Hsieh & Urquiola, 2002).

Public and private-voucher schools compete to enroll as many students as they can. The Ministry of Education prescribes a national curriculum, establishes set standards for school infrastructure, and supervises school operations for public, private-voucher schools. On the other hand, the public and private-voucher schools have the freedom to make their own decisions, operate schools, and establish their own school policies such as innovation in the curriculum and other areas that they consider will make them more attractive to their customers (Gauri, 1998).

Accountability System

One of the central components of the Chilean educational system is the notion of free choice, in which parents would be able to obtain more efficient educational services so long as they had the possibility of choosing between schools, that schools, in return, would compete for students (Fontaine & Eyzaguirre, 2001). Thus the information provided by standardized test is intended to inform parents so they are able to make knowledgeable decisions when choosing schools for their children.

Chile is one of the pioneers in Latin America in terms of its incorporation of standardized tests as a measurement of the country's quality of education, which is classified as one of the best in the region (Crouch, 2005; Manzi, 2014; Meckes, 2007). The current system of evaluation of quality of education was put into place in 1988, when the Chilean government implemented the *Sistema Nacional de Evaluación de Calidad de la Educación* (Educational Quality Measurement System, SIMCE), which is used to measure student achievement. In 1990, an education law (Ley Orgánica Constitucional de Enseñanza, LOCE) provided the legal foundation and operational details of the SIMCE, currently operated by the Ministry of Education, which is responsible for its development, implementation, and diffusion (Taut, Cortés, Sebastian, & Preiss, 2009). The SIMCE test has two main objectives; one is to obtain information about whether or not students are achieving the minimum standard of learning established by the central government and to determine the quality of education that schools are providing (Fontaine & Eyzaguirre, 2001).

From low-stakes to high-stakes accountability systems

Chile has been collecting and disseminating student performance data for more than 20 years, the purpose of this has been to provide information to parents about school quality. In fact, until 2007 the SIMCE test scores had created little in the way of

incentives or consequences for teachers or school administrators (Mizala & Urquiola, 2007). However, in 2008 the Chilean congress approved the implementation of the *Subvención Escolar Preferencial* (preferential voucher program) also known as SEP. This law increased the voucher for those students of low socioeconomic status, and it also gave additional resources to schools in proportion to the number of low-income students enrolled. Public and private-voucher schools participate voluntarily in this program, but in so doing they commit to improving certain academic and management goals established and reinforced by the Chilean government (Mizala & Torche, 2013) Even though SEP is voluntary, in 2011, after 3 years of its implementation, 99 percent of the public schools and 73 percent of the private-voucher schools participated in SEP program (Mineduc, 2012). In order to evaluate whether or not schools were achieving the goals in their contracts, the SEP program ranks schools according to students' performance on standardized tests, the socioeconomic status of the student body, and other indicators of the quality of education. If schools do not accomplish the benchmarks established by the SEP program, schools could be sanctioned and in the worst case, closed (Elacqua, Santos, Martinez, & Urbina, 2011). The Chilean government had invested around $1,300 million dollars on this program (Mineduc, 2012).

With the implementation of the SEP program, Chile has started a new type of educational reform centered on the use of mechanisms of pressure and accountability (Carrasco & San Martin, 2012). In fact, in 2015 the Chilean government will implement a nationwide high-stakes accountability system approved in 2011 by the Chilean congress (Mineduc, 2011), in which all schools will be required to participate. In summary, this study analyzes the main elements of the market-oriented policies in the Chilean education. I study the effect of the introduction of competition and student achievement as well as explore the teachers' responses to the system—academic and non-academic

strategies — during the period before and after the implementation of the SEP program. This information is fundamental for policy-makers as the implementation of the nationwide high-stakes accountability system in 2015 approaches and the educational reform that the current government is planning to implement.

STRUCTURE, METHODS AND SUMMARY OF FINDINGS

This dissertation is divided into six chapters. Chapter 1, the introduction, is followed by the Chapter 2, the literature review which introduces the central ideas and assumptions concerning competition, free choice, and accountability in education and the main problems that researchers face in evaluating the impact of these reforms.

Chapter 3, "The Winner Takes it All: Competition and Student Achievement," relates to the first objective of this research, and explains the two main approaches to understanding the macro level effects of competition in education. While the neoliberal perspective suggests that the whole system would benefit from an increase of competition, critical theory argues that the potential benefits of competition would be "canceled out" by the negative effect of an increase in socioeconomic and class segregation.

This chapter asks whether competition has an impact on student achievement and whether or not the impacts of competition are uniform across the country. In order to answer these questions I analyze data from the *Sistema Nacional de Evaluación de Calidad de la Educación* (Educational Quality Measurement System, SIMCE) to measure student achievement in reading and math. The SIMCE test is administered to all students in the country, in all of the educational establishments, including public, private-voucher and private-paid schools. The SIMCE test is administrated every year in 4th grade and every two years in 8th grade. Information about students' grades is provided by the

Chilean Ministry of Education, from a database called *Rendimiento por Estudiante.* The data encompasses the individual level and informs on the student's final grade obtained, rate of attendance, and whether or not the student passed or failed his/her grade.

I use ordinary least square (OLS) models with fixed effects at the municipality level as a way to address the endogeneity problem of competition and test scores. I use the percentage of students enrolled in 4th grade in private-voucher schools per municipality as a measurement of competition. The results of the regression models show that competition has a negative effect on student achievement at the national level. For example, for a one percent increase in competition, student achievement decreases by 0.06 points on the reading test and 0.08 points on the math test. However, the coefficient of competition for the metropolitan region, where the capital city is localized, is positive and statistically significant, but relatively small compared with the standard of deviation of the SIMCE test, which is 50 points. For example, a one percent increase in competition implies only a 0.15 point increase on the reading test and a 0.1 point increase on the math test in 4^{th} grade. In summary, this chapter offers empirical evidence that runs contrary to assumptions made by advocates of market-oriented reform in education, those who hypothesized that the breakdown of the government's educational monopoly and the introduction of competition and free choice in the system would improve the overall quality of the system in the long run. Instead, the introduction of the private-voucher schools only reproduces the differences already existing in the country.

In Chapter 4, "Under Pressure: Teachers' Responses to the High-stakes Accountability Program," corresponds to the second objective of this dissertation, which is to understand how teachers respond to the accountability system. It asks whether teachers have increased the use of academic strategies as a way to improve students' performance on standardized tests and what difficulties and constraints the Chilean

educational system imposes on teachers and school administrators in order to increase the use of academic strategies under the high-stakes accountability system. The chapter draws on the categorization developed by Daniel Koretz (2008) regarding how teachers respond to high-stakes accountability systems. In particular, it explores whether or not teachers who work in schools that voluntarily enroll in the high-stakes program have increased the use of academic strategies after the implementation of the program. The data analyzed comes from the *Cuestionario Docentes* (teacher questionnaire). This questionnaire is administrated by the Ministry of Education and is answered by each teacher whose students will be evaluated by the SIMCE tests, which inquires about his or her professional training and the contents that were taught during the school year, among others. The majority of the questionnaires were answered by one teacher who taught the three classes tested by SIMCE: math, reading, and sciences.

I use descriptive statistics to analyze teacher questionnaires for teachers of 4th and 8th grade before and after the implementation of the high-stakes program in 2008 to find if they are using academic strategies to improve student learning. Furthermore, I complement the findings from the teacher questionnaires with the content analysis of semi-structured interviews that I conducted in Chile with 23 academics, policy-makers and teachers. The results of this chapter indicate that schools that participate in the high-stakes accountability systems have not developed a set of academic strategies, such as working harder, working more or finding better methods of teaching to improve student learning. I argue that this is because the Chilean government is pushing schools to produce outcomes in the short term with measurable results, and academic strategies do not necessarily offer these results.

Chapter 5, "The Invisible Children: What Standardized Tests Left Behind" connects to the third objective of this dissertation and focuses on the unintended

responses of teachers under the high-stakes accountability system. I use Campbell's law to explore the use of non-academic strategies after the implementation of the high-stakes program. In particular, this chapter investigates whether or not low-achieving students are more likely to be excluded from the test taking pool and how this could impact in schools' ranking. I use the data from the Educational Quality Measurement System (SIMCE) for 2nd, 4th, and 8th grades from 2002 to 2013 and parent questionnaire. The parent questionnaire is a survey that all families need to answer if their children take the SIMCE test. It provides information about family income, the education level of mother and father, among other variables and it is administrated by the Ministry of Education.

I use a logistic regression with fixed effects at the school level to explore whether or not low-achieving students are more likely to be excluded from the test-taking pool after the implementation of the high-stakes accountability system in 2008. In order to estimate how the schools' rankings based on standardized tests change if the students excluded from the test-taking pool would be able to take the standardized test. I use imputation models on data from the 2nd, 4th and 8th grades to impute the missing values of those students who did not take the standardized tests and I use the values imputed, previously excluded, to recalculate the classification of schools and explore whether or not it is possible to observe any difference including those students who did not take the standardized test.

The main findings of this chapter show that after the implementation of the high-stakes accountability system in 2008, students who obtained a grade of B- or lower were 27.3 percent more likely to have not taken the reading test and in math 35.7 percent more likely. Using multiple imputed data, I have recalculated schools' rankings by using the imputed values of those students who did not take the SIMCE test, and the number of

schools classified as “insufficient” increases as well as those classified as “intermediate.” The number of schools ranked as “advanced” decreases in 4th and 8th

Chapter 6 concludes the main findings of each chapter, providing the policy implication of the results in the context the educational reform propose by the current government. Finally, limitations of this study and future research in this area are presented.

References

Betebenner, D. W., Howe, K. R., & Foster, S. S. (2005). On School Choice and Test-Based Accountability. *Education Policy Analysis Archives*, *13*(41), 22.

Campbell, D. (1979). Assessing the impact of planned social change. *Evaluation and Program Planning*, *2*(1), 67–90.

Campbell, D. E., West, M. R., & Peterson, P. E. (2005). Participation in a national, means-tested school voucher program. *Journal of Policy Analysis and Management*, *24*(3), 523–541. Retrieved from http://dx.doi.org/10.1002/pam.20113

Carnoy, M., & McEwan, P. J. (2000). The Effectiveness and Efficiency of Private Schools in Chile ' s Voucher System. *Educational Evaluation and Policy Analysis*, *22*(3), 213–239.

Carrasco, A., & San Martin, E. (2012). *Clasificacion de Escuelas en la Nueva Institucionalidad Educativa: Contribucion de Modelos de Valor Agregado para una Responsabilidad Justa* (No. 53). Santiago, Chile.

Correa, J., Parro, D., Reyes, L., & Ugarte, G. (2012). *El impacto del voucher diferenciado en los resultados académicos: el caso chileno*.

Crouch, L. (2005). Political Economy, Incentives, and Teacher's Union: Case Studies in Chile and Peru. In E. Vegas (Ed.), *Incentives to Improve Teaching: Lessons from Latin America* (pp. 389–417). Washington, D.C.: The World Bank.

Elacqua, G., Santos, H., Martinez, M., & Urbina, D. (2011). *¿Estamos preparados para cerrar las malas escuelas en Chile? Impacto sobre equidad en el acceso a educación de calidad*. Santiago, Chile. Retrieved from http://www.comunidadescolar.cl/documentacion/FONIDE/Informe Final-Gregory Elacqua-UDP-511083.pdf

Falabella, A. (2013). *Accountability Policy Effects Within School Markets A Study in Three Chilean Municipalities*. University of London. Retrieved from http://www.academia.edu/12249727/ACCOUNTABILITY_POLICY_EFFECTS_WITHIN_SCHOOL_MARKETS_A_STUDY_IN_THREE_CHILEAN_MUNICIPALITIES

Fontaine, L., & Eyzaguirre, B. (2001). Una Estructura que Presione a las Escuelas a Hacerlo Bien *. In *¿Qué Hacer Ahora? Propuestas para el Desarrollo.* (pp. 66–111). Santiago, Chile: Centro de Estudios Publicos.

Gauri, V. (1998). *School Choice in Chile: Two Decades of Educational Reform*. Pittsburgh: University of Pittsburgh Press.

Haertel, E., & Herman, J. (2005). *A Historical Perspective on Validity Arguments for Accountability Testing* (Vol. 1522). Loas Angeles, CA.

Hamilton, L. S., Stecher, B. M., & Klein, S. P. (2002). *Test-Based Accountability in Education*. Santa Monica, CA: RAND Ecation.

Hsieh, C., & Urquiola, M. (2002). *When schools compete , how do they compete ?* (No. 43). New York City.

Hursh, D. (2007). Assessing No Child Left Behind and the Rise of Neoliberal Education Policies. *American Educational Research Journal*, *44*(3), 493–518.

Kremer, M., & Sarychev, A. (2000). *Why Do Governments Operate Schools?* Retrieved from http://post.economics.harvard.edu/faculty/kremer/papers.html.

Ladd, H. F. (2002). School Vouchers: A Critical View. *Journal of Economic Perspectives*, *16*(4), 3–24.

Levin, H. (1998). Educational vouchers_ Effectiveness, choice, and costs.pdf. *Journal of Policy Analysis and Management*, *17*(3), 373–392.

Manzi, J. (2014, March 17). ¿Ganamos algo eliminando el Simce? *La Tercera*, pp. 2–3. Santiago, Chile. Retrieved from www.latercera.cl

Meckes, L. (2007). Evaluación y Estándares: Logros y desafíos para incrementar el impacto en calidad educativa. *Revista Pensamiento Educativo*, *40*(1), 351–371.

Mineduc. Ley 20.529: Sistema Nacional de Aseguramineto de la Calidad de la Educacion Parvularia, Basica y Media y su Fiscalizacion (2011). Valparaiso. Retrieved from http://www.cned.cl/public/secciones/seccionsnac/normativa/LEY-20529_27-AGO-2011.pdf

Mineduc. (2012). *Impacto de la Ley SEP en SIMCE: Una Mirada a 4 años de su Implementacion*. Santiago, Chile. Retrieved from http://www.mineduc.cl/usuarios/bmineduc/doc/201208301544260.EVIDENCIASCEM8.pdf

Mizala, A., & Romaguera, P. (2000). School Performance and Choice The Chilean Experience. *The Journal of Human Resources*, *35*(2), 392–417.

Mizala, A., & Torche, F. (2013). ¿Logra la subvención escolar preferencial igualar los resultados educativos? *Espacio Publico*, *09*, 1–36. Retrieved from http://www.espaciopublico.cl/publicaciones/detalle.tpl?id=19

Mizala, A., & Urquiola, M. (2007). *School Markets: The Impact of Information Approximating Schools' Effectiveness* (No. 13676). Cambridge, MA. Retrieved from http://www.nber.org/papers/w13676

Nechyba, T. J. (2000). Mobility, Targeting, and Private-School Vouchers. *The American Economic Review*, *90*(1), 130–146. Retrieved from http://www.jstor.org/stable/117284

Perticara, M., Román, M., & Selman, J. (2013). *Rendimiento escolar y uso de los recursos de la ley SEP* (No. I-297). Retrieved from http://fen.uahurtado.cl/wp-content/uploads/2010/07/I-2971.pdf

Taut, S., Cortés, F., Sebastian, C., & Preiss, D. (2009). Evaluating school and parent reports of the national student achievement testing system (SIMCE) in Chile: access, comprehension, and use. *Evaluation and Program Planning*, *32*(2), 129–37. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/19026445

Torche, F. (2005). Privatization Reform and Inequality of Educational Opportunity : *Sociology of Education*, *78*(Mmi), 316–343.

Villarroel, G. (2012). *Mejoramiento en Resultados Académicos de la Educación Básica en Chile: ¿Primeros efectos de la ley de Subvención Escolar Preferencial (SEP)?* Universidad de Chile.

Wall, D. (2000). The impact of high-stakes testing on teaching and learning : can this be predicted or controlled? *System*, *28*, 499–509.

West, B. (1997). Education Vouchers in Principal and Practice: A Survey. *The World Bank Research Observe*, *12*(1), 83–103.

Chapter 2: Literature Review

A generalized concern among both developed and developing countries is the quality of education, and how to improve it (Chapman, Wesley, & Jr, 2000; Figlio & Getzler, 2002). One common solution implemented in many English-speaking Western nations, such as England, North Ireland and Wales in the late 1980s and in the early 1990s in the United States has been the introduction of market–oriented policies into the educational arena (Betebenner et al., 2005; Hursh, 2007).

Of late, this topic has received more attention in Latin America because at least 12 countries have introduced diverse school choice programs and accountability regimes into their education systems – reforms largely based on models championed by policy makers in the United States. One recent example is the educational reform passed in 2012 in Mexico, which included the creation of a national agency overseeing evaluations and plans to increase the accountability of teachers (Andere, 2012). Policy-makers assume that competition among schools and the free choice of parents could improve the quality of education, while standardized tests are viewed as a way to provide a "quality indicator" to the consumer (parents) and an "objective assessment" of student achievement within the educational market (Hursh, 2007; Robertson, 2000). This approach combines both "free choice" and "accountability" as complements to one another. In this sense, test-based accountability systems could help parents to choose by providing information about the quality of educational services, and school choice facilitates accountability by promoting competition for enrollment (Betebenner et al., 2005).

Diverse studies have shown that test-based accountability systems might create incentives for teachers and school administrators to use non-academic strategies, such as

coaching students to take standardized tests, or excluding low-performing students from the test-taking pool, as a way to improve schools' performance in standardized tests (Koretz, 2008b). However, it is hard to evaluate how extensive these negative effects are and how they could impact schools and students (Stecher & Barron, 2001). In the case of Chile, some studies have analyzed the use of non-academic strategies under the Chilean high-stakes accountability system. These researches have used a descriptive approach, however, and it is therefore not possible to know how extensive and significant the use of these strategies is. In this sense, the present research is a contribution to the current literature as it not only quantifies the extension of the use of non-academic strategies over time, but it also determines how significant of an impact these strategies have on schools' ranking.

In the next section, I will introduce the central assumptions behind competition, free choice, and accountability in education and the main problems that researchers face in evaluating the impact of these reforms.

MARKET-ORIENTED POLICIES IN EDUCATION: COMPETITION, FREE CHOICE, AND ACCOUNTABILITY IN EDUCATION

In his book *Capitalism and Freedom* (1962), Milton Friedman examines the contemporary debate about the role of markets and governments in education. Friedman assumed that the private sector is superior to the public sector in terms of providing quality and efficiency, and therefore argued that fostering competition among schools, could improve the quality of education (Friedman, 1962). Friedman proposed a type of "educational marketplace," not only to provide for choice, especially among poor students without the option to attend a private school, but also to create incentives to improve efficiency in the delivery of educational services and innovation in education (Levin, 2002). In the "educational marketplace" model of education, resources for

schools are not guaranteed, and as such schools compete for financial resources while parents choose the schools that provide a better education for their children. As a result, schools with poorer performance would lose students and eventually disappear from the market. For example, a large-scale voucher plan might "permit competition to develop," thus leading to the development and improvement of all schools" (Friedman, 1962). Friedman suggests that most funding should be directed to students themselves through a voucher mechanism: “The subsidization of institutions rather than people has led to an indiscriminate subsidization of all activities appropriate for such institutions, rather than of activities appropriate for the state to subsidize” (Friedman, 1962). He argued that the change from institutional to individual funding would increase the competition between schools and promote quality, efficiency, differentiation, and innovation among educational institutions.

In an educational system, which allows to parents to choose the school for their children will create competition, incentives for innovations, and differentiation among schools, accountability, and an overall improvement of the system. The assumption is that individuals are the basic maximizing units, and therefore an educational system that provides options such as public (magnet schools) and private school (voucher/charter) would imply maximized actions for parents and students. A system designed to offer options has a direct effect on customers (parents and students) because it allows them to maximize their own goals. Also, when individuals pursue a particular objective, the collective good likewise benefits, because if every member of a certain group is better off, then the entire community is better off as well. Proponents argue that efficiency and equity can only be achieve by an increase in “free choice” (Betebenner et al., 2005; Robertson, 2000) as parents choose the best school for their children based on academic performance, values, and the quality of the schools. They are trying to maximize the

return on their investment in the long run, calculating the cost and benefits of choosing certain types of schools. Parents, particularly those who are in the most disadvantaged socioeconomic group, would thus benefit from the creation of "competition" among schools and "free choice" (Betts & Loveless, 2005).

As the neoliberal approach has gained popularity in the educational arena, test-based accountability systems have been seen as a necessary condition for improving the quality of education. In recent decades, policy efforts in the United States have been focused on implementing these types of programs (Chapman et al., 2000; Cullen & Reback, 2006; Figlio & Getzler, 2002; Hamilton et al., 2002; E. Hanushek & Raymond, 2005; Jacob, 2007; Kim, 2004; Neal & Schanzenbach, 2010; Stecher & Barron, 2001). The assumption of those who advocate for the implementation of a standardized test-based accountability system is that the mechanism of rewards and sanctions created by test-based accountability systems will encourage teachers and school administrators to become better at their jobs and maximize students learning, as a result. In line with this perspective, raising expectations and standards should promote better opportunities for students and narrow the academic achievement gap between low and high-performing students (Haertel & Herman, 2005).

PROBLEMS TO EVALUATE MARKET-ORIENTED POLICY IN EDUCATION

This section presents the main difficulties that researchers face in testing the hypothetical benefits of competition, free choice, and test-based accountability in education.

Competition: Omitted variable bias and endogeneity issues

One of the fundamental arguments in support of school choice is that increased competition will lead to an improvement in the quality of the education (Dee, 1998;

Hsieh & Urquiola, 2003) However, testing this hypothesis is difficult due to omitted variable bias and issues of endogeneity.

The problem of omitted variables occurs when factors that confound the relationship between public and private-voucher schools are omitted from the analysis. For example, the equilibrium demand for private schools is significantly affected by several dimensions of socioeconomic status (Gemello & Osman, 1984; Lankford & Wyckoff, 1992) , and it is also well known that the socioeconomic characteristics of families are related to students' academic achievement (E. A. Hanushek, 1986; Haveman & Wolfe, 1995). Another source of omitted variables could come from "ability bias," where private-voucher schools select the best students by way of a selection test or, in effect, by implementing higher tuition rates, a practice that reduces the average ability and educational outcomes in public schools (Belfield & Levin, 2002a). If an empirical model does not adequately incorporate omitted variables (explained above), the conclusions drawn between competition and quality of education could lead to misleading inferences (Dee, 1998).

Endogeneity issues arise when, in a school district where public schools provide low quality of education, the demand for private-voucher schools rise and create a negative relationship between public school quality and private schooling enrollment (Dee, 1998). This is a problem if researchers try to estimate the effects of competition and quality of education in school districts with bad public schools, mainly because the low quality of education creates incentives to offer alternatives such as voucher schools, and it can appear that voucher schools cause this low quality of education, instead of the other way around (C. Hoxby, 2003)

Additional, problems arise with the way that competition is defined. Typically, competition is measured using the Herfindahl Index (HI), which is defined as the sum of

the squares of per-unit enrollments over total enrollments, bounded between 0 (full competition) and 1 (monopoly) (Borland & Howsen, 1992). Most empirical studies that use the HI as a continuous variable have found that competition has only weak or null effects on students' performance. On the other hand, studies that have used the HI to classify school districts according to high or low levels of competition have found more statistically significant results (Belfield & Levin, 2002a). Another measure of competition is the private schooling enrollment share, which is the percentage of students enrolled in private-voucher schools in a given school district (Dee, 1998; E. A. Hanushek & Rivkin, 2003; C. Hoxby, 1994)

The endogeneity problem involving competition and quality and the omitted variable bias has different implications. First, the existence of simultaneity suggests that Ordinary Least Squares could underestimate the effect of competition on students' academic achievement because families in school districts with poor-quality public schools are more likely to request private-voucher schools as an alternative, and private-voucher schools may draw "good" students away from public schools. Second, the existence of omitted variable bias implies that OLS could overestimate the true effect of competition on students' academic achievement due to the fact that this relationship is influenced by important, but unobserved, determinants of student achievement (Dee, 1998).

One of the fundamental arguments in favor of private-voucher schools is the idea that public schools will become more effective due to increased competition, but as noted above, testing the effects of competition has proven to give elusive results (Carnoy & McEwan, 2003).

Free choice: Selection bias and innovation

Proponents of market-oriented reform argue that quality education can only be achieved if parents are able to choose freely among the schools available to their children. The assumption is that parents would choose the best schools for their children, and therefore "bad" schools (poor quality, low achievement, etc.) would close because they will not have enough students to keep the school running and the overall quality of the educational system would improve as a result of parental choice (Betts & Loveless, 2005) However, also it is hypothesized that school choice will exacerbate educational inequalities among schools because students' academic achievement is influenced by their peers as well as the students' socioeconomic status (Kremer & Sarychev, 2000; Levin, 1998)

The implementation of a large-scale private-voucher program is likely to increase the racial and socioeconomic stratification of schools because many parents use the social and ethnic composition of a school to judge school quality. Moreover, low-income families are in a less favorable position to choose the best school for their children than higher-income families because low-income families have less access to information and they cannot pay for transportation when private-voucher schools are located far from the poorest neighborhoods. Moreover, because private-voucher schools are allowed to select their students, they will most likely pick students of medium or high socioeconomic status, as they are easier and less costly to educate (Ladd, 2002).

As a consequence, comparing public vs. private-voucher school students could be unfair and lead to biased conclusions because students are not fully comparable (Carnoy, 2001; Levin, 1998). The findings on students' achievement from public vs. private-voucher schools will be contaminated by selection bias. Students in private-voucher schools may come from higher socioeconomic statuses and may have unobserved

characteristics—abilities, motivated families, etc.—that could increase the likelihood of the student obtaining a higher score on achievement test, regardless of the schools they attend (Greene, Peterson, & Du, 1998).

Under the Chilean school choice system, families who are of medium or high socioeconomic status may be more likely to take advantage of the free choice than those who come from the lowest socioeconomic strata of society. This is mainly because wealthier families have better access to information, greater ability to afford transportation, and more experience with choices and alternatives (Levin, 1998). It has been argued that private-voucher schools would lead families with the most resources, and thus a greater ability to improve the public schools, to exit the public system and opt for private-voucher schools, rather than exercise their voice and create pressure that could lead to the improvement of public education (D. E. Campbell et al., 2005; Hirschman, 1970). Furthermore, unlike public schools, private-voucher schools can select their students through selection tests and they are able to decide to increase their tuition depending according to their own interests and priorities. Therefore, vouchers only cover a certain percentage of the tuition, and parents need to pay extra money to register their children in these schools. Thus, private-voucher schools can improve the average ability level of their student body, and create "brain drain" from public schools that would lead to and increase existing segregation based on the socioeconomic status of the students (Hsieh & Urquiola, 2003; Levin, 1998; P. J. McEwan, 2000a; Neal, 2011).

Another argument that advocates of "free choice" employ claims that under a system that allows parents to choose schools for their children, schools should also increase the diversity of education models, such as curricula and educational innovation and, therefore, further increase the choices available to parents. However, Rawls (2001) points out that even if private-voucher schools are not found to be a state actor, in order

to avoid the government's bureaucracy, they will still be subject to state regulations through the acceptance of the voucher package. For example, every sovereign has an obvious interest in education; this interest is intensified when education is funded by taxpayer money in the form of a voucher. Therefore, schools that accept vouchers are not sufficiently private to maintain independence from state influence, and they must accept the regulations and restrictions imposed by the government. In brief, it is almost impossible for private-voucher schools to exist outside the constraints of government rules while maintaining the freedom of allowing individuals to choose and implement their own education system (Rawls, 2001).

Furthermore, Brown (1992) argues that schools provide services that can be divided into two categories. Primary services are those that affect the labor market characteristic of students, and include instruction in traditional subjects such as mathematics, biology, and history, as well as workplace socialization such as vocational training, athletics activities, music, etc. Secondary services can be provided outside of school or not provided at all without having an effect on the students' labor market options. The most common secondary service is related to religious instruction. All schools provide a full range of primary services, and private and public schools try to differentiate themselves in the services supplied in order to be more attractive for parents or students. However, private-voucher schools have difficulties finding an empty niche in the schooling market, except by differentiating themselves on a secondary dimension. In fact, they tend to look like public schools, quite independent of the institutional form. Brown points out that risk aversion among clients (parents) and the ability of the schools to mitigate that uncertainty by providing equal opportunity are powerful forces that serve to make schools look the same.

In fact, private-voucher and public schools organize themselves in a way that allows parents to make value judgments about the quality of education that their children will receive in these institutions. This is particularly the case with families who are concerned with the future of their children and with how education can improve their opportunities in the labor market. Parents take as few risks as possible, as they know little about the impact of one curriculum or another, of one teaching method or another; they will be conservative about the school they choose for their children. If parents observe that a school differs radically from the traditional ones, it increases the risk that parents will make the wrong decision about the best school for their children. Therefore, public and private non-voucher schools need to adapt according to their customers' requirements, and parents will push the schools that their children attend to behave like other schools (Brown, 1992).

At the earliest grade levels, schools know little about the students' abilities. Therefore, we can observe the greatest similarities in what is taught and how it is taught to different students as a market response to uncertainty concerning student ability at this point. An example of this is a school's use of whole classroom instruction and significant levels of uniformity in curriculum. Rather than creating innovations, private-voucher schools tend to be more conservative and closer to traditional schools than Friedman and other scholars have suggested (Benveniste, Carnoy, & Rothstein, 2003).

Accountability: Non-behavioral and behavioral issues.

The difficulties in evaluating the impact of the accountability system in education can be categorized as non-behavioral and behavioral. The non-behavioral issues are related to test reliability and even though large-scale tests are a relatively objective way to measure the most valued aspect of students achievement, they are neither perfect nor

comprehensive measures (Hout & Stuart, 2011). On the other hand, behavioral issues are those that relate to how behavioral responses to testing modify or affect test validity (Koretz, 2008a)

Non-Behavioral Issues

One of the main concerns in psychometrics is related to measurement error, which in the case of standardized tests is focused on inconsistencies across different measurements of a single student, due to a variation from the items that being tested or from fluctuation in a given individual's performance over time (Koretz, 2008a). For example, standardized test are constructed using certain type of questions and are focused on specific contents. As a consequence, the tests only measure a subset of the area being tested (sample), and test scores will differ from one period to another, even though the test is taken by the same individual (Hout & Stuart, 2011). The way that standardized tests are reported also creates fluctuations in the students' or schools' performance. For example, Linn (2007) analyzed The National Assessment of Educational Progress (NAEP) data on the achievement gap between minority and majority groups in order to show how relevant these distortions can be. In 4th grade and 8th grade, the gap between white versus Hispanics and black students decreased from 1996 to 2005. However, by using the "percent above basic" and "percent above proficient" statistics, the author demonstrated contradictory results: some comparisons indicate the gap between groups reduced over time, while others appeared to show an upward trend (Linn, 2007).

Finally, test-based accountability systems require choosing a metric that will be used to summarize the results. Different metrics commonly categorize schools very differently; for example, students' performance can be compared using "cohort to cohort models" and "valued added models." Cohort to cohort is the principal model used in the

United States (Koretz, 2008a), and in this model students' performance are analyzed by successive cohorts in a particular grade, to observe if the students' performance in that grade is improving over time (Hout & Stuart, 2011). However, each cohort is not necessarily comparable to another, and therefore some variations in the test scores may be explained by the characteristics of students in each cohort, rather than by teachers or schools. In order to address this problem, value-added models have been developed. This type of model analyzes the growth paths of students' performance using longitudinal data that includes different observations of each individual student over time (Raudenbush, 2004). Value-added models try to control for differences across individuals in order to quantify the percentage of student growth that could be accurately explained by schools or teachers (Hout & Stuart, 2011). This type of model seems to be fairer, as it compares the effectiveness of different schools and teachers, but studies have shown that curricular differentiation becomes a serious issue with this type of model. For example, in middle school mathematics, a standardized test that has been designed for general use will most likely offer weak coverage of focus areas for teachers of high achieving students. As a consequence, valued-added models may incorrectly categorize these teachers as "inefficient" (Mccaffrey, Lockwood, Koretz, & Hamilton, 2003)

Behavioral Issues

This category refers to how teachers or school administrator responses to testing could affect test validity. For example, if a particular school shows an improvement in standardized tests, it is difficult to know what teachers or schools have done in order to increase the students' performance. There are different strategies that teachers and schools can use to improve students' achievement; including the implementation of test-based accountability systems. For example, teachers could seek out better teaching

methods or they could spend more time with their students, which can lead to improved student learning and therefore an improvement in their performance on standardized tests. However, there are other strategies that also can improve student achievement, but they are not the type of skills or actions that are encouraged by the implementation of test-based accountability systems. These practices include "teaching to the test," which is the practice of coaching students with similar questions and settings as are present in the real standardized test. Teachers and school administrators could also identify low-performing students and simply exclude them from the test-taking pool.

Koretz (2008) presents seven typologies to describe the responses of teacher to high-stakes accountability systems.

Table 2-1: Response of Teachers to High-Stakes Accountability System

Teacher responses	Actions	Outcomes	Strategy
(1) Working More Effectively	Finding better methods of teaching	Good	Academic
(2) Teaching More	Spending more time overall	Good	Academic
(3) Working Harder	Giving more HW or harder assignments	Good	Academic
(4) Reallocation	Shifting resources, including time, to emphasize the subjects and type of questions on the test	Ambiguous	Ambiguous
(5) Alignment	Matching the curriculum more closely to the material covered on the test	Ambiguous	Ambiguous
(6) Coaching Students	Preparing students using old tests	Bad	Non Academic
(7) Cheating	Exclusion of students, classification of students with special needs, disciplinary sanctions	Bad	Non Academic

Source: Measuring Up: What Educational Testing Really Tells Us, Daniel Koretz (2008)

As Table 2-1 shows, in the first group of responses, teachers work more effectively, teach more, and work harder to produce meaningful gains in test scores. That means, these three strategies could produce higher levels of achievement on substantive areas of the test, which suggest that students have improved in the intended area (Koretz & McCaffrey, 2001).

In the second group of responses, where teachers shift resources (reallocation) or match their curriculum more closely to the material covered on the test (alienation), the results are ambiguous. This is because if the test is a good measurement of the contents taught in the grade's curriculum, students could benefit from these types of strategies. But if the test does not appropriately address the main aspects of the curriculum, the students' learning could be negatively affected by these types of strategies (Kamenetz, 2014) Studies have found that teachers have few incentives to place attention on those students who are already proficient or that have a lower chance of passing the standardized test (Neal & Schanzenbach, 2010) and, therefore, pay more attention and spend more time on those students who could likely pass the standardized test (Booher-Jennings, 2005; Dworkin, 2005; Hursh, 2007). Furthermore, schools and teachers could respond to the pressure to raise test scores by decreasing time spent on subjects that are not part of or are less emphasized on the exams. Thus, students could spend more time working on math or reading while reducing the time allotted to the sciences, history, arts, foreign languages, or writing because they are not topics that will be tested. This can lead to the narrowing and oversimplification of the curriculum (E. Baker et al., 2010; Carnoy, Loeb, & Smith, 2000; Hamilton et al., 2002; Stecher & Barron, 2001; Winters, Trivitt, & Greene, 2010).

Alignment between curriculum and standardized test might be seen as a practice that protects against score inflation, because it would focus teachers' instruction on the

areas and elements that are relevant and valuable for those who designed the test-based accountability system. As a consequence, because the teacher's instruction and test content are coherent, assuming that the test scores provide a good measure of students' learning, an increase in students' performance is meaningful, not inflated. However, another point of view argues that the definition of a particular set of standards is arbitrary, as well as the claim that student learning can be extracted from the tests, is uncertain (Koretz & McCaffrey, 2001).

The third group of teachers' responses to the high-stakes accountability system has negative consequences on student learning, and these are not the skills or strategies that policy makers try to incentivize in schools through the implementation of test-based accountability systems. Coaching or "teaching to the test" refers to many types of tests preparation. Coaching can be focused on elements of the test that are entirely non-substantive. For example, students can take preparation tests every week in the same format as the original test, or teachers can figure out that the standardized test always includes regular quadrilaterals and triangles in area and perimeters problems, and therefore teachers can make the decision do not teach about irregular polygons or figures with more than four sides in his or her class anymore (Koretz, 2003). Cheating implies many types of actions such as providing answers to students, correcting students' responses, allowing them use external material during testing, violating test administration procedures, or excluding low-performing student from the test-taking pool, among others. However, cheating, unlike coaching, cannot increase the validity of scores (Koretz & McCaffrey, 2001).

The problems associated with cheating are not in and of themselves an argument against the test-based accountability system. However, they do indicate a need for the careful interpretation of gains in standardized tests, because incentives can create a

variety of responses from schools or teachers, and some of them are clearly illegitimate and do not improve the students' learning (Koretz, 2002).

INCENTIVES IN TEST-BASED ACCOUNTABILITY SYSTEMS

Test-based accountability systems are based on the assumption that if a particular outcome or result is rewarded, this outcome will increase the likelihood of being observed because it is going to be reinforced by the incentive. However, the difficulty is that teachers' performance measures are necessarily focused on the areas of the job that are more easily quantified, while they neglect the qualitative aspects of the job that are more difficult to measure (Hout & Stuart, 2011). For example, if surgeons are penalized for the mortality rates of their patients, they will accept cases with less risk of death; if a company rewards its workers for the company's earnings, the workers might manipulate the company's earnings report in order to gain the rewards (Rothstein, 2008).

In the educational arena, schools are responsible for diverse aspects of the students' education: civic participation, cognitive, emotional and physical development, among others. Furthermore, under a test-based accountability system, schools are responsible for the students' performance and tasked with ensuring that all students either pass the minimum standards or meet very high ones. In this context, where schools need to meet different and diverse goals for the students' education and there are limited resources to do so, schools are obligated to make difficult tradeoffs among these multiple goals (Dixit, 2002). Thus, economic theory points out that when complex jobs involve many tasks and workers' performance is evaluated based on one objective measure, workers most likely shift their effort among the various tasks to be focused on the most easily observable aspects of a multi-dimensional task. Thus, the measurement by itself will create incentives to narrow multiple objectives of the job to those that are actually

measured, disregarding the overall mission of the organization (G. Baker, 1992; Feltham & Xie, 1994; Holmstrom & Milgron, 1991). According to Baker, Jensen and Murphy (1998) one of the major disadvantages of measuring a job with multi-dimensional tasks, using an objective system, is that employees will "game the system" by optimizing their performance on the task that is evaluated or measured (G. Baker, Jensen, & Murphy, 1988).

From a psychological point of view, studies in this area have shown that the external rewards have mixed responses among individuals (Deci, Ryan, & Koestner, 1999). For example, receiving a reward could have the significant effect of enhancing subsequent motivation, although sometimes it seems to have the effect of diminishing subsequent motivation, and in cases it appears to have no effect. Studies have shown that the differences are related to the degree to which individuals perceive the rewards as external pressure, or if it is something under his or her control. If they think the reward limits their autonomy because they do not have control over it, the result of this incentive tends to be negative. However, if the reward is not perceived as an external pressure, but instead as something that simply signifies competence, this type of reward will most likely have a positive effect (Harackiewicz, Abrahams, & Wageman, 1987).

Madaus & Clarke argue that test-based accountability systems could create incentives that result in diverse and even unintended negative consequences (Madaus & Clarke, 2001). Social psychologist, Donald Campbell, corroborates this with his eponymous law to explain the same phenomena: When social decision-making is focused on a single quantitative indicator, he suggests that "the more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social process it is intended to monitor" (D. Campbell, 1979). In this sense, test-based accountability systems in

education provide an excellent example of Campbell's law, as standardized tests are essentially the sole indicator used to evaluate quality of education and they carry heavy consequences for schools. Teachers and school administrators could, therefore, be incentivized to use non-academics strategies to ensure that the test scores give them favorable results (Brian & Levitt, 2003; Nichols & Berliner, 2007). Even though the studies are not conclusive, the overall tendency shows that test-based accountability programs, which are focused on high-stakes testing, might have important negative consequences for schools and students (Volante, 2007).

Non-academic strategies include the exclusion of low achieving students from the test-pool and their reclassification as special education students, suspending them or encouraging them to stay home during the testing window (Booher-Jennings, 2005; E. Hanushek & Raymond, 2005; Kane & Staiger, 2002). In the test-based accountability system, schools and teachers face great pressure to raise test scores and therefore it is logical for schools to leave some students, especially those who do not perform well, out of the test-taking pool. And rather than improving the overall achievement of all students, which is the main objective of the implementation of test-based accountability systems, schools and teachers may simply find diverse strategies that allow them to exclude the "weak" students from the test-taking pool (Hursh, 2007). If students do not take reading, math or science tests, the student score is reported as a missing value and, therefore, this student's score does not count in the calculation the school's average. In some cases, if a student does not take any of these standardized tests his or her scores will not be reported as a missing value; instead, this student could be excluded from the dataset. Figlio (2003) shows that during the testing window, low-achieving students are most likely to receive longer suspensions than high-performing students who have committed similar violations, and students who received long suspensions are much more likely to miss the

standardized test (Figlio, 2003). Moreover, studies have shown that schools classify low-achieving students as special education in order to exclude them from the test-taking pool (Figlio & Getzler, 2002; Haney, 2000; Jacob, 2005).

This study will test Campbell's law in the context of test-based accountability systems in Chile, and will explore whether or not it is possible to observe some degree of alteration or distortion before and after the implementation of high-stakes accountability systems in Chile. Numerous studies have shown that high-stakes testing could artificially inflate students' academic achievement. Particularly when the stakes are too high, teachers and schools could have incentives to act fraudulently or to cheat (Heckman, Heinrich, & Smith, 2002; Kane & Staiger, 2002) In fact, the findings indicate that school administrators and teachers are inclined to create the appearance of the improvement of students' scores on standardized tests by using a collection of non-academic practices and, as such, gaming the system (Dworkin, 2005; Wall, 2000). However, there are no studies in Chile that quantify the extension and the effect of the use of non-academic strategies on students' performance.

References

Andere, E. (2012). *The 2012 Education Reform in Mexico: From Intention to Action: The Expert Take*. Retrieved from https://mexicoinstitute.wordpress.com/2012/12/19/the-2012-education-reform-in-mexico-from-intention-to-action-the-expert-take/

Baker, E., Barton, P., Darling-Hammond, L., Haertel, E., Ladd, H., Linn, R., … Shepard, L. (2010). *Problems with the Use of Students Test Scores to Evaluate Teachers* (No. 278) (Vol. 03). Washington, DC.

Baker, G. (1992). Contracts Performance Measurement. *The Journal of Political Economy*, *100*(3), 598–614.

Baker, G., Jensen, M., & Murphy, K. (1988). Compensation and Incentives: Practice vs. Theory. *The Journal of Finance*, *XLIII*(3), 593–616.

Belfield, C. R., & Levin, H. M. (2002). The Effects of Competition Between Schools on Educational Outcomes: A Review for the United States. *Review of Educational Research*, *72*(2), 279–341. doi:10.3102/00346543072002279

Benveniste, L., Carnoy, M., & Rothstein, R. (2003). *All Else Equal: Are Public and Private Schools Different?* NYC: RoutledgeFalmer.

Betebenner, D. W., Howe, K. R., & Foster, S. S. (2005). On School Choice and Test-Based Accountability. *Education Policy Analysis Archives*, *13*(41), 22.

Betts, J., & Loveless, T. (2005). *Getting choice right: ensuring equity and efficiency in education policy*. Washington, D.C.: Brookings Institution Press.

Booher-Jennings, J. (2005). Below the Bubble: "Educational Triage" and the Texas Accountability System. *American Educational Research Journal*, *42*(2), 231–268.

Borland, M. V, & Howsen, R. M. (1992). Student academic achievement and the degree of market concentration in education. *Economics of Education Review*, *11*(1), 31–39. Retrieved from http://www.sciencedirect.com/science/article/pii/027277579290019Y

Brian, J., & Levitt, S. (2003). Rotten Apples: An investigation of the Prevalence and Predictors of Teacher Cheating. *The Quaterly Journal of Economics*, (August), 843–877.

Brown, B. (1992). Why governments run schools. *Economics of Education Review*, *11*(4), 287–300. doi:10.1016/0272-7757(92)90038-5

Campbell, D. (1979). Assessing the impact of planned social change. *Evaluation and Program Planning*, *2*(1), 67–90.

Campbell, D. E., West, M. R., & Peterson, P. E. (2005). Participation in a national, means-tested school voucher program. *Journal of Policy Analysis and Management*, *24*(3), 523–541. Retrieved from http://dx.doi.org/10.1002/pam.20113

Carnoy, M. (2001). *School Vouchers: Examining the Evidence*. Washington, D.C.: Economic Policy Institute.

Carnoy, M., Loeb, S., & Smith, T. L. (2000). *Do Higher State Test Scores in Texas Make For Better High School Outcomes? The Civil Right s Project*.

Carnoy, M., & McEwan, P. (2003). *Choosing Choice: School Choice in International Perspective*. (G. Plank, David & Sykes, Ed.)*Comparative Education Review* (Vol. Vol. 42, N). New York: Teachers College Press.

Chapman, D. W., Wesley, C., & Jr, S. (2000). Can High Stakes National Testing Improve Instruction: Reexamining Conventional Wisdom. *International Journal of Education Development*, *20*, 457–474.

Cullen, B. J., & Reback, R. (2006). *Tinkering Toward Accolades: Schools Gaming Under a Performance Accountability System* (No. 12286). Retrieved from www.nber.or/papers/w12286

Deci, E., Ryan, R., & Koestner, R. (1999). A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation. *Psychological Bulletin*, *125*(6).

Dee, T. S. (1998). Competition and the quality of public schools. *Economics of Education Review*, *17*(4), 419–427. doi:10.1016/S0272-7757(97)00040-X

Dixit, A. (2002). Incentives and Organizations in the Public Sector An Interpretative. *Journal of Human Resources*, *37*(4), 696–727.

Dworkin, a. G. (2005). The No Child Left Behind Act: Accountability, High-Stakes Testing, and Roles for Sociologists. *Sociology of Education*, *78*(2), 170–174.

Feltham, G. A., & Xie, J. (1994). Performance Measurement Congruity and Diversity in Congruity Principal / Agent Relations. *The Accounting Review*, *69*(3), 429–453.

Figlio, D. (2003). *Testing, Crimen and Punishment* (No. 11194). New York City. Retrieved from http://www.nber.org/papers/w11194

Figlio, D., & Getzler, L. (2002). *Accountability, Ability and Disability: Gaming The System* (No. 9307). Cambridge, MA.

Friedman, M. (1962). *Capitalism and Freedom*. Chicago: University of Chicago Press.

Gemello, J. M., & Osman, J. W. (1984). Estimating Private the Demand Enrollment for. *American Journal of Education*, *92*(3), 262–279. Retrieved from http://www.jstor.org/stable/1085005

Greene, J., Peterson, P., & Du, J. (1998). School Choice in Milwaukee. In *Learning from School Choice* (pp. 329–344). Brookings Institution Press.

Haertel, E., & Herman, J. (2005). *A Historical Perspective on Validity Arguments for Accountability Testing* (Vol. 1522). Loas Angeles, CA.

Hamilton, L. S., Stecher, B. M., & Klein, S. P. (2002). *Test-Based Accountability in Education*. Santa Monica, CA: RAND Ecation.

Haney, W. (2000). The Myth of the Texas Miracle in Education. *Education Analysis Policy Archives*, *8*(41). Retrieved from http://epaa.asu.edu/ojs/article/view/432

Hanushek, E. A. (1986). The Economics of Schooling: Production and Efficiency in Public Schools. *Journal of Economic Literature*, *24*(3), 1141–1177 CR – Copyright © 1986 American Econ. doi:10.2307/2725865

Hanushek, E. A., & Rivkin, S. G. (2003). The Economics of School Choice. In C. M. Hoxby (Ed.), (Vol. I).

Hanushek, E., & Raymond, M. (2005). Does school accountability lead to improved student performance? *Journal of Policy Analysis and Management*, *24*(2), 297–327.

Harackiewicz, J. M., Abrahams, S., & Wageman, R. (1987). Performance evaluation and intrinsic motivation: The effects of evaluative focus, rewards, and achievement orientation. *Journal of Personality and Social Psychology*, *53*(6), 1015–1023. doi:10.1037//0022-3514.53.6.1015

Haveman, R., & Wolfe, B. (1995). The Determinants of Children's Attainments: A Review of Methods and Findings. *Journal of Economic Literature*, *33*(4), 1829–1878 CR – Copyright © 1995 American Econ. doi:10.2307/2729315

Heckman, J. J., Heinrich, C., & Smith, J. (2002). *The Performance of Performance Standards* (No. 9002). Cambridge, MA.

Hirschman, A. O. (1970). *Exit, voice, and loyalty: responses to decline in firms, organizations, and states*. Cambridge, Mass: Harvard University Press, 1970.

Holmstrom, B., & Milgron, P. (1991). Multitask Principal-Agent Analyses : Incentive Contracts , Asset Ownership , and Job Design. *Journal of Law, Economics, & Organizations*, *7*(January 1991), 24–52.

Hout, M., & Stuart, E. (2011). *Incetives and Test-Based Accountability in Education*. Retrieved from http://standardizedtests.procon.org/sourcefiles/incentives-and-test-based-accountability-in-education.pdf

Hoxby, C. (1994). *Do Private Schools Provide Competition For Public Schools?* (No. 4978). Retrieved from http://www.nber.org/papers/w4978

Hoxby, C. (2003). School choice and school competition: Evidence from the United States. *Swedish Economic Policy Review*, *10*, 9–65.

Hsieh, C., & Urquiola, M. (2002). *When Schools Compete, How Do They Compete? An Assessment Of Chile's Nationwide School Voucher Program*. *NBER Working Paper Series* (Vol. 10008). doi:10.1016/j.jpubeco.2005.11.002

Hursh, D. (2007). Assessing No Child Left Behind and the Rise of Neoliberal Education Policies. *American Educational Research Journal*, *44*(3), 493–518.

Jacob, B. (2005). Accountability, incentives and behavior: the impact of high-stakes testing in the Chicago Public Schools. *Journal of Public Economics*, *89*(5-6), 761–796.

Jacob, B. (2007). *Test-Based Accountability and Students Achievement: An Investigation of Differential Performance on NAEP and State Assessments* (No. 12817). Cambridge, MA.

Kamenetz, A. (2014). When Teachers, Not Students, Do The Cheating. Retrieved from http://www.npr.org/blogs/ed/2014/09/29/352494603/when-teachers-not-students-do-the-cheating

Kane, T. J., & Staiger, D. O. (2002). The Promise and Pitfalls of Using Imprecise School Accountability Measures. *Journal of Economic Perspectives*, *16*(4), 91–114.

Kim, S. (2004). Accountability and School Reform in the U . S . Public School System. *KEDI Journal of Educational Policy*, *1*(1), 2–35.

Koretz, D. (2002). Limitations in the Use of Achievement Tests as Measures of Educators ' Productivity. *The Journal of Human Resources*, *37*(4), 752–777.

Koretz, D. (2003). Using Multiple Measures to Address Perverse Incentives and Score Inflation. *Educational Measurement: Issues and Practice*, *22*(2), 18–26.

Koretz, D. (2008a). Further Steps Toward the Development of an Accountability-Oriented Science of Measurement. In K. Ryan & L. Shepard (Eds.), *The Future of Test-Based Educational Accountability* (pp. 71–91). Taylor and Francias, Hoboken.

Koretz, D. (2008b). *Measuring Up: What Educational Testing Really Tells Us*. Harvard University Press.

Koretz, D., & McCaffrey, D. (2001). *Toward a Framework for Validating Gains Under High-Stakes Conditions* (Vol. 1522). Los Angeles, CA.

Kremer, M., & Sarychev, A. (2000). *Why Do Governments Operate Schools?* Retrieved from http://post.economics.harvard.edu/faculty/kremer/papers.html.

Ladd, H. F. (2002). School Vouchers: A Critical View. *Journal of Economic Perspectives*, *16*(4), 3–24.

Lankford, H., & Wyckoff, J. (1992). Primary and secondary school choice among public and religious alternatives. *Economics of Education Review*, *11*(4), 317–337. doi:10.1016/0272-7757(92)90040-A

Levin, H. (1998). Educational vouchers_ Effectiveness, choice, and costs.pdf. *Journal of Policy Analysis and Management*, *17*(3), 373–392.

Levin, H. (2002). A Comprehensive Framework for Evaluating Educational Vouchers. *Educational Evaluation and Policy Analysis*, *24*(3), 159–174.

Linn, R. L. (2007). *Considerations when selecting and combining measures. Paper presented at a workshop at the National Academies*. Washington, DC.

Madaus, G., & Clarke, M. (2001). The adverse impact of high stakes testing on minority students: evidence from 100 years of test data. In G. Orfield & M. Kornhaber (Eds.), *Raising standars or raising barriers? Inequality and high stakes testing in public education* (p. 51). The Century Foundation. Retrieved from http://files.eric.ed.gov/fulltext/ED450183.pdf

Mccaffrey, D., Lockwood, J. R., Koretz, D., & Hamilton, L. (2003). *Evaluating Models for Teacher Accountability*. Retrieved from http://www.cgp.upenn.edu/pdf/rand.pdf

McEwan, P. J. (2000). *The Potential Impact of Large-Scale Voucher Programs*. *Review of Educational Research* (Vol. 70). Retrieved from http://rer.sagepub.com/cgi/doi/10.3102/00346543070002103

Neal, D. (2011). How Vouchers Could Change the Market for Education. *The Journal of Economic Perpectives*, *16*(4), 25–44.

Neal, D., & Schanzenbach, D. (2010). Left behind by design: proficiency counts and test-based accountability. *The Review of Economics and Statistics*, *92*(May), 263–283.

Nichols, S. L., & Berliner, D. C. (2007). Collateral damage: how high-stakes testing corrupts America's schools. Cambridge, Mass.: Harvard Education Press. Retrieved from http://catalog.hathitrust.org/Record/005555685

Raudenbush, S. W. (2004). *Schooling, Statistics and Poverty: Can We Measure School Improvement?* New Jersy. Retrieved from http://files.eric.ed.gov/fulltext/ED486417.pdf

Rawls, A. (2001). Eliminating Options Through Choice. *Emory Law Journal*, *50*(1), 363–396.

Robertson, S. (2000). *A Class Act: Changing Teachers' Work, Globalization and the State*. New York, NY 10001: Routledge.

Rothstein, R. (2008). *Holding Accountability to Account: How Scholarship and Experience in Other Fields Inform Exploration ofPerformance Incentives in Education* (No. 04). Retrieved from http://s4.epi.org/files/2014/holding-accountability-to-account.pdf

Stecher, B. M., & Barron, S. (2001). Unintended Consequences of Test-Based Accountability When Testing in " Milepost " Grades. *Educational Assessment*, *7*(4), 259–281.

Volante, L. (2007). *Evaluating Test-Based Accountability System: An International Perspective*. Stockholm, Sweden.

Wall, D. (2000). The impact of high-stakes testing on teaching and learning : can this be predicted or controlled? *System*, *28*, 499–509.

Winters, M. a., Trivitt, J. R., & Greene, J. P. (2010). The impact of high-stakes testing on student proficiency in low-stakes subjects: Evidence from Florida's elementary science exam. *Economics of Education Review*, *29*(1), 138–146.

Chapter 3

The Winner Takes it All: Competition and Student Achievement

There are two approaches that explain the macro-level effect of the implementation of large-scale private-voucher programs in education. The first one comes from a neoliberal perspective and suggests that whole system would benefit from increased competition in education, as schools would have incentives to improve their performance in order to attract students and, therefore, the quality of education would dramatically improve if parents were allowed to choose freely between schools (Hsieh & Urquiola, 2003). A second approach comes from critical theory and argues that the potential benefits of competition would be "cancelled out" by the negative effect of increasing socioeconomic and class segregation among students that would be promoted by school choice programs (Bellei, 2007).

From a neoliberal approach, in the "educational marketplace" model of education, resources for schools are not guaranteed, and as such schools compete for financial resources, parents choose the schools that provide a better education for their children. As a result, schools with poorer performance would lose students and eventually disappear from the market. For example, a large-scale voucher plan might "permit competition to develop… and improvement of all schools" (Friedman, 1962). On the other hand, critical theory suggests that the implementation of a large-scale private-voucher program will not have the expected effect mainly because private-voucher schools would most likely select students from a medium or high socioeconomic status because they are easier and less costly to educate – this effect is called "cream skimming" (Ladd, 2002). The evidence suggests that the effects of cream skimming for targeted programs are relative small, but

they could pose a serious problem for universal voucher programs (Altonji, Huang, & Taber, 2005).

The effect of increased choice on many of the relevant empirical issues surrounding school choice programs are still controversial (Nechyba, 2003). Most researchers in Chile have focused on the relationship between privatization reform and social segregation (Bellei, 2012; Elacqua, 2006; Flores, 2008; F. Torche, 2005; Valenzuela, Bellei, & De Los Rios, 2008). The impact of competition on student achievement in Chile has received less attention in the literature; in fact, there are no articles published in peer-reviewed journals about this issue.

One of the main problems in establishing a causal relationship between competition and student achievement comes from the idea that the degree of competition in a school district could be endogenous to student performance (Ahlin, 2003). That is, if a school district provides a low quality of education, private-voucher schools could have incentives to open and compete with public schools. Using a traditional, ordinary least squares model, researchers could erroneously conclude that an increase in competition would decrease students' achievement. Previous studies that explore the relationship between competition and student achievement in Chile have tried to avoid the endogeneity problem by using instrumental variables on two-stage least square models (Auguste & Valenzuela, 2004; F. A. Gallego, 2002; Hsieh & Urquiola, 2003). However, the instrumental variables such as the numbers of priests per municipality, urbanization rates, population per municipality, and the distance to the closest big city have been criticized as valid instruments (Bellei, 2007). Furthermore, the period used by these studies is limited. For example, Auguste & Valenzuela and Gallego only used one year of observation, while Hsieh & Urquiola's data is not comparable and includes less than half of the municipalities in Chile.

In this sense, this study makes a contribution to the previous literature in Chile in two ways. First, I improve the quality and the period of observations used in previous studies, as I use longitudinal data at student level from the 4th grade from the whole country during the period of 2002-2013, provided by The National Assessment of Mathematics and Spanish Achievement, known as the *Sistema Nacional de Evaluación de Calidad de la Educación* (Educational Quality Measurement System, SIMCE). Second, I used a fixed effects model at the municipal level to deal with the endogeneity problem that previous studies tried to solve using an instrumental variable. This is possible because the data that I have collected allows me to compare the same municipality through time for more than 10 years.

In this chapter, I explore two hypotheses that advocates of neoliberal policies have argued in support of competition in education. Namely, that the introduction of private-voucher schools will increase competition in the system and that as a result of that, the quality of education will improve. Moreover, I test the assumption that competition would benefit the whole educational system, especially low-income families or minority groups, as they will be able to exercise their free choice to select the best schools for their children (Friedman, 1962). I explore these hypotheses guided by the following research questions: how does competition affect students' achievement? Does competition have similar impacts on the performance of students who are educated in the capital city, Santiago, and outside of it? In the next section I explain the theoretical foundation behind the introduction of competition education and the main effect that competition has had on student achievement and social segregation in Chile.

THE MACRO-LEVEL EFFECT OF PRIVATE SECTOR EXPANSION IN EDUCATION

A central argument used by advocates of school choice programs is that the competition that is generated between schools would improve quality of education under the assumption that the private sector is superior to the public sector in terms of providing quality and efficiency (Friedman, 1962). While resources for the students are not guaranteed, schools must compete for financial resources, and parents choose those schools that provide a better education for their children. As a result, those schools with poor performance would lose students and eventually disappear from the market (Flores, 2008).

The school choice system assumes that competition among schools to obtain student vouchers will develop strong incentives not only to satisfy student and parent requirements, but also to improve educational productivity, reducing dropout rates, increasing enrollment, etc. In fact, an educational marketplace is not only creating a choice, especially among poor students without the option to attend a private school, but also creating incentives to improve efficiency in the delivery of educational services and innovation across the system. Competition would lead to a greater range of choices and increased efficiency and innovation in education, as schools have financial incentives to attract and retain their enrolled students ((Levin, 2002). In summary, a large-scale voucher plan might "permit competition to develop," thus leading to the "development and improvement of all schools" (Friedman, 1962).

On the other hand, opponents of voucher school programs argue that the implementation of a large-scale private-voucher program is likely to increase the racial and socioeconomic stratification of schools because a lot of parents use the social and ethnic composition of a school to judge school quality (Ladd, 2002).

Moreover, low-income families are in a less favorable position to choose the best school for their children than higher income families, given that low-income families have less access to information and that they cannot pay for transportation when private-voucher schools are located far from the poorest neighborhoods. Moreover, if private-voucher schools are allowed to select their students, they will most likely pick students from a medium or high socioeconomic status because they are easier and less costly to educate (Ladd, 2002).

The Impact of School Competition in Chile

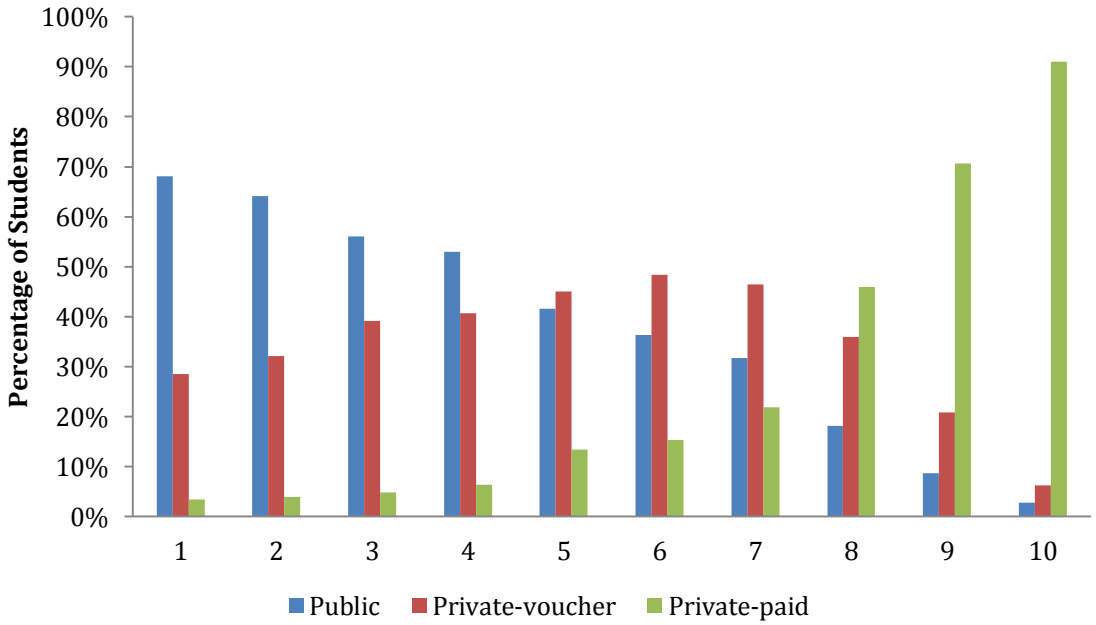
The empirical evidence available suggests that there is not a clear and unique answer as to the impact of the expansion of the private-voucher sector in Chile. Some claim that if competition has been found to have an impact on student achievement, it is irrelevant and not widely observed at a national level. At the same time, diverse studies suggest that the introduction of choice and competition into the system could cause increased socioeconomic segregation and an achievement gap among students (Bellei, 2012)

One widely accepted hypothesis that explains the relationship between competition, segregation, and the achievement gap is called the "cream skimming" effect. This suggests that the potential benefits of competition would be "canceled out" by the negative effects of increasing socioeconomic and class segregation among students promoted by school choice programs. While public schools need to accept all types of students, private-voucher schools might have different mechanisms by which they select their student body, a process that increases the likelihood of enrolling students with good academic performance. As a result, students with higher socioeconomic status benefit from the "peer effect" of attending private-voucher schools, as these are sites with higher

accumulation of economic, social, and cultural capital. Instead, students of low socioeconomic status are concentrated in public schools, segregated and socially isolated. The existing evidence suggests that the cream skimming effect for targeted programs is relatively small, but could be a serious problem for universal voucher programs (Altonji et al., 2005).

In the case of Chile, it is possible to observe the "cream skimming" effect, due to the fact that under the educational system implemented in 1981, public schools cannot select students through admission tests or interviews. This is because public schools have a "public function," and cannot discriminate against students based on test performance or socioeconomic status. Private-voucher schools, however, can select students according to their own criteria. Therefore, the private-voucher schools select their students from the advantaged (middle class). These students are "easier" to educate and their families contribute their own financial and political resources to the private-voucher schools. Historically, the wealthiest group has been educated in private non-voucher schools that receive no government funding, while the lower–middle and middle-income sectors are educated in private-voucher schools, and the poorest sector in public schools. In fact, while students from low-income families are more likely to attend public schools, students from a medium or medium-high socioeconomic status tend to enroll in private-voucher schools, and those students from high socioeconomic status are most likely to attend private schools (Elacqua, 2006; González, Mizala, & Romaguera, 2004). For a visual representation of this data, the next figure shows the distribution of types of schools attended by income decile in 2000. As is evident in the figure 3-1, different income deciles attend different types of school, as explained above.

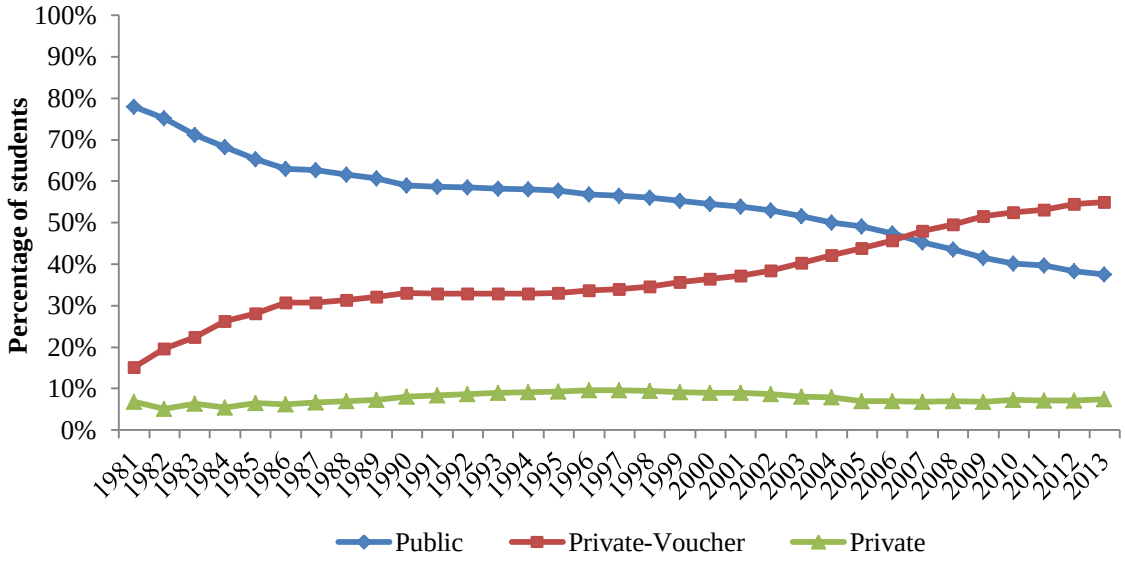
Figure 3-1: Type of Schools Attended by Income Decile; 2000



Source: González, Mizala & Romaguera's calculations (2004) based on the 2000 CASEN survey (Chilean Ministry of Planning)

In the case of competition, which is defined as the number of students enrolled in primary and secondary private-voucher schools, the field was relatively stable from 1990 to 2000. But since 2000, schools faced a stronger degree of competition and, as a result of that, the number of students enrolled in private-voucher schools surpassed public schools in 2013. As Figure 3-2 illustrates, almost 78 percent of Chilean students attended public schools in 1981, and 15 percent attended private-voucher schools. However, after more than 30 years of privatization reform, public school enrollment decreased to 37.5 percent, with 55 percent of Chilean students attending private-voucher schools in 2013.

Figure 3-2: Primary and Secondary Enrollment from 1981 to 2013



Source: Author's calculations based on Estadísticas Educacionales, Ministry of Education Chile (1990 - 2013).

It is important to clarify that private-voucher schools can be divided into two groups: for-profit and non-profit schools. Non-profit schools are part of religious institutions (usually Catholic or Evangelical churches) and for-profit schools belong to a private owner or anonymous society. The growth of private-voucher school from the 90s can be explained mainly by for-profit schools. In fact, for-profit private-voucher schools grew 81 percent from 1990 to 2008. In contrast, the number of non-profit private-voucher schools only increased by 16 percent during the same period (Corvalan, Elacqua, & Salazar, 2009). The growth of the per capita income of families, demographic changes, and the different rules of financing between public and private-voucher schools could explain the growth of private-voucher schools since the 90s (Paredes & Pinto, 2009). Particularly, in 1993 the Chilean government created a new regulation about the way that schools could be financed; it is called *financiamiento compartido* (shared funding). This new system allowed private-voucher schools to decide whether to increase their tuition

according to their own interests and priorities. Therefore, vouchers only cover a certain percentage of the tuition, and parents need to pay extra money to register their children at these schools. In contrast, public schools only can use "shared funding" on high school education and with more restrictions than private-voucher schools. The new finance system created conditions for the proliferation the private-voucher schools, particularly those that were for-profit (Corvalan et al., 2009; Paredes & Pinto, 2009).

How does this increase in competition affect student achievement? In order to establish a causal relationship between competition and student achievement, one must first address a major problem: the degree of competition in any school district could be endogenous to student performance (Ahlin, 2003). As I explained in Chapter 2, the existence of simultaneity implies that Ordinary Least Square (OLS) may underestimate or overestimate the effect of private-voucher schools on student achievement in public schools because high quality public schools can draw students away from private-voucher schools (Dee, 1998). As a result, studies that use private schooling to measure competition typically use two-stage least square (2SLS) or instrumental variable (IV) approaches (Belfield & Levin, 2002b). Some studies have used the extent of scale decentralization as a proxy for competition (Hoxby, 2000; Urquiola, 2000) whereas others have used the concentration of Catholics or number of Catholic priests in school districts as an instrumental variable (Dee, 1998). Finally, other researchers have employed a combination of variables, for example, the degree of privatization of elderly care, the degree of privatization of infrastructure, the degree of privatization of social services, the average population distance from schools, and the average cost per pupil (Ahlin, 2003).

The evidence on the impact of competition on student achievement in Chile shows the following findings using ordinary least square (OLS) and panel data at the

school level from more than 15 years from the 1980s to the mid-1990s, Carnoy and McEwan (1999) use private enrollment as a measure of inter-school competition, and they found that an increase in competition has a small, but positive effect on the SIMCE test given to lower middle-class students who attend one large group of public schools in Santiago. However, for the rest of the schools located in the capital city, the effect on of increased competition was not statistically significant, and for the rest of Chile, the impact of increasing competition on test scores for students who attend public schools was slightly negative, but statistically significant. Finally, the study did not find significant results on the overall test scores for public and private-voucher schools in those school districts where the number of students enrolled in private-voucher schools had increased more rapidly (P. McEwan & Carnoy, 1999).

Hsieh and Urquiola (2003) use a two-stage least square model (2SLS) and panel data for about 150 municipalities, at the school level, to evaluate the relationship between competition and student achievement, competition is defined as the proportion of students who attend private school at the municipal level. In order to address problems of endogeneity, the researchers use three instrumental variables. The first two variables were urbanization rates and the population of a municipality in 1982, and the third instrument is the inter-quartile range in years of schooling, observed among adults. The study found no evidence that an increase in competition among private-voucher schools and public schools had improved student academic performance, declined repetition rates, and increased years of schooling. Hsieh and Urquiola found that the average SIMCE test score did not rise any faster in municipalities where the private-voucher schools made greater inroads, and the grade repetition average and grades by age measures had worsened in these communities. The overall findings of this study would suggest that school choice did not improve educational outcomes in Chile (Hsieh &

Urquiola, 2003). However, these findings have be criticized because the information available before the introduction of private-voucher schools is very limited, and it is difficult to establish a causal relationship using this data (Bellei, 2007).

Similarly, Gallego (2004) using the number of Catholic priests by municipality as an instrumental variable in a two-stage least square model, looked at 280 municipalities (across 345 municipalities) and the metropolitan region (where the capital city is located) as proxies for local educational markets and where competition is defined as the share of total enrollment in private school in each municipality. The capital city (Santiago) was excluded from the study, student performance was measured by the SIMCE test in 2002 (for 4th grade), and the unit of analysis is at the student level. The study found that Chilean students have received significant benefits from competition, but the effect of competition on student performance is smaller for students who attend public school than for those who attend private-voucher schools (F. A. Gallego, 2004). Finally, the use of the number of Catholic priests per municipality has been criticized as valid instrument because religious schools (non-profit) only represent a small proportion of the private-voucher schools (one third) and religious congregations are more involved with private non-voucher schools (Bellei, 2007).

Using a similar approach to that of Gallego, in order to avoid endogeneity problems and selection bias, Auguste and Valenzuela (2004) used a two-stage least square model, but with different types of instrumental variables: the municipality's population, as a measure of market size, and the distance to the closest big city, as a measure of entry costs as instrumental variables. They analyzed data from 8th grade classrooms in 2000, at the student level, but the data was aggregated at the municipal level, with the exception of the capital city, which pooled together several neighboring municipalities. They defined competition as the private sector market share, that is, the

ratio of private-voucher and private-paid school enrollment to total enrollment at the municipal level. The authors concluded that competition has had a positive but moderate effect on student achievement, an increase of one standard deviation in competition is linked with an increase for 0.4 standard deviations on average test scores at the municipality level, and also the increase of competition was associated with higher levels of segregation (Auguste & Valenzuela, 2004). Table 3-1 summarizes the main studies conducted in Chile on the relationship between competition and student achievement.

Table 3-1: Previous Studies on Competition and Students Achievement in Chile

Year	Authors	Unit of analysis	Model	Grades	Period	Type
1999	Carnoy & McEwan	school	OLS	4th grade	1987-1995	Unpublished paper
2003	Hsieh & Urquiola	school	2SLS	4th grade	1981-1998	Working paper
2004	Gallegos	student	2SLS	4th grade	2002	Working paper
2004	Auguste & Valenzuela	student	2SLS	8th grade	2000	Dissertation

In summary, the evidence available in Chile does not allow for conclusive results as to the systematic effects of the implementation and expansion of private-voucher schools (Bellei, 2007).

Competition and Geographical Location of Schools

The level of competition varies across municipalities, where municipality size and socioeconomic characteristics are central in explaining the effect of competition on students achievement (Auguste & Valenzuela, 2004). This study uses this finding to explore the different effects of the introduction of private-voucher schools between the wealthiest region (Metropolitan) versus the poorest region (Araucanía) in Chile,

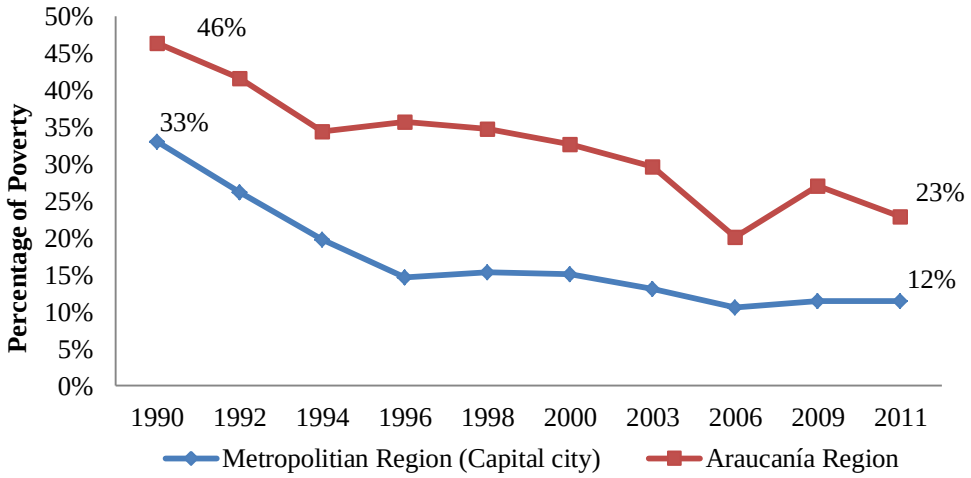
exploring whether or not the effects of competition on students achievement are dissimilar, according to the conditions in which schools are located.

The metropolitan region, where the capital city is located, is the wealthiest region in Chile, and it is also home to the main private industries and state institutions, such as universities, enterprises, services, hospitals, and government institutions. The Metropolitan region's contribution to GDP is the highest among Chile's 13 regions. For example, in 2012 its contribution was 44.4 percent (Banco-Central, 2015). The Metropolitan region also presents the highest household income level (Casen, 2013). In the educational arena, the illiteracy rate is 1.8 percent in the Metropolitan region, versus the national average of 3.3 percent Also, the average years of schooling in this region is higher than the national average, 11.3 years for the Metropolitan region versus 10.6 years at the national level (Casen, 2011b). In contrast, the contribution to national GDP for the Araucanía region was only 2 percent in 2012 (Banco-Central, 2015). The illiteracy rate is more than 2 percentage points above the national average, 5.4 percent for the Araucanía versus 3.3 percent at the national level. While the years of schooling in this region is lower than the national average – 9.3 years for the Araucanía region versus 10.6 years at the national level (Casen, 2011a).

Figure 3-3 shows the evolution of poverty in the Metropolitan and Araucanía region from 1990 to 2011. Araucanía region is the poorest region in Chile, with 23 percent of its population living below the poverty threshold. The national average poverty level is 14.4 percent. Instead, the Metropolitan region has 12 percent of its population living in poverty, which is below the national average. Furthermore, the Araucanía region is not only the poorest region, but a high percentage of its population is also indigenous and rural (Casen, 2009). In terms of development indicators, an official report from the

Chilean government estimates that the Araucanía region is more or less 20 years behind of the rest of Chile (Gobierno de Chile, 2011).

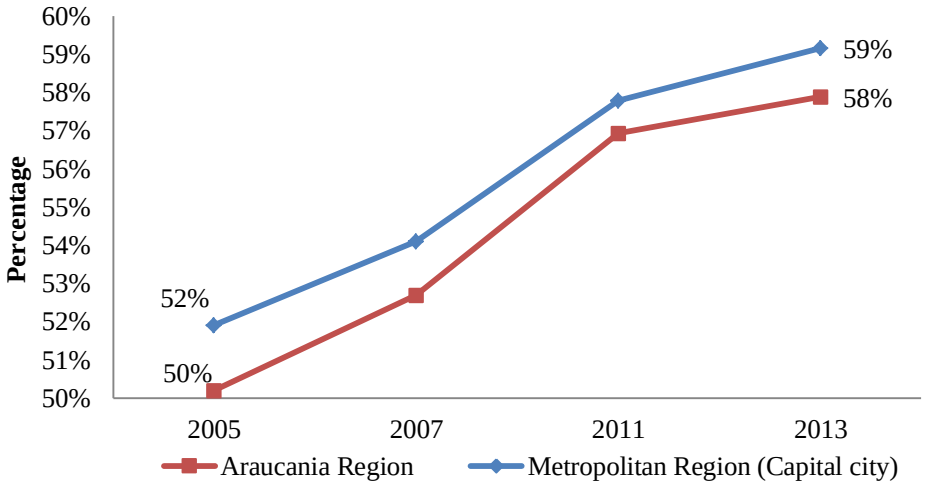
Figure 3-3: Percentage of Poverty in Metropolitan and Araucanía Regions from 1990 to 2011



Source: Author's calculation based on Casen 1990-2013

Even though there are clear differences in terms of human development indexes between the Metropolitan and Araucanía region, as I explained above, both regions share the same pattern in terms of the participation of private-voucher schools, as Figure 3-4 shows – the percent of students enrolled in private-voucher schools increased 7 percent from 2005 to 2013 in the Metropolitan region and 8 percent in the Araucanía region over the same period.

Figure 3-4: Percentage of Primary Education Students Enrolled in Private-Voucher Schools in Metropolitan and Araucanía Regions from 2005 to 2013



Source: Author's calculations using the dataset "Matricula por estudiante" Centro de Estudios (2015)

I use these differences in terms of human development and similarities in relation to the participation of private-voucher schools between regions to analyze the impact of competition on student achievement. I hypothesized that competition could have different outcomes depending of the conditions where it is implemented. For example, since the Metropolitan region has comparative advantages in terms of human and social capital over other regions in Chile, I argue that in these conditions the introduction of competition among schools could have a positive impact on student achievement. Instead, when schools compete the Araucanía region, which is 20 years behind the rest of Chile in terms of human development indexes, I hypothesized that in disadvantaged conditions, competition can be negative and may have unintended consequences on student achievement.

This is relevant from a public policy prospective, as the same policy could have a different impact based on the place that it is implemented and, therefore, national policies need to take into account regional differences in order to ensure to achieve the expected goals.

METHODS AND DATA

I use two different data sets on the Chilean educational system. Since the data come from different sources, the following section explains the information and sources contained in the data set in more detail:

(a) Sistema Nacional de Evaluación de Calidad de la Educación (Educational Quality Measurement System, SIMCE). The SIMCE test is administered to all students in the country, in all of the educational establishments, including public, private-voucher and private non-voucher schools. Testing takes place in the fourth, eighth, and twelfth grades, in different years (Mizala & Romaguera, 2000), and since 2012 it has been applied yearly to the second grade. SIMCE scores report the performance of schools in different subject areas (math, reading, and sciences) in comparison to previous years, other schools, and other grade levels within the same school. This research relies on data from the standardized SIMCE test applied to 4th grades from 2002 to 2013 (see Table 3-2); the unit of analysis is student-level. SIMCE not only measures the academic performance of students, but also the socioeconomic characteristics of their families and teachers (Contreras, Bustos, & Sepulveda, 2010).

The variables included in the SIMCE test have three levels: 1) the student level, which includes information such as math and verbal scores, gender, and whether or not the student attended preschool; 2) parent questionnaire, that includes individual family characteristics such as family type, the educational level of the student's parents, the income of the family, etc. and; 3) the school level, which includes information such as private-voucher, public or private school, the location, and the socioeconomic status of the students who attend this school. The data set from students and parents are merged using a unique student ID, and the schools' datasets are merged with student and parent data using the unique ID of the school.

Table 3-2: Grades and Years When SIMCE Test Was Administered

Grades	Years									
	2002	2005	2006	2007	2008	2009	2010	2011	2012	2013
4^{th}	√	√	√	√	√	√	√	√	√	√
# cases	269,863	254,287	247,090	242,919	238,903	223,908	232,107	219,239	239,243	226,005

b) The second data source comes from the Centro de Estudios (Department of Research of Chilean Ministry of Education) and is called *rendimiento del estudiante* (student achievement). The data encompasses individual-level achievement and offers the final grade obtained by a student, their rate of attendance, and whether or not the student passed or failed during the period between 2002 and 2012. Furthermore, it provides a student identification number that can be used to merge with data from the National Assessment of Mathematics and Spanish Achievement (SIMCE).

Problems of endogeneity arise in an attempt to estimate the impact of competition on student achievement (Dee, 1998), if researchers try to estimate the effects of competition and quality of education in a municipality with bad public schools, mainly because low quality of education creates incentives to offer alternatives such as private-voucher schools, it can appear that private-voucher schools cause low quality education, instead of the other way around (C. Hoxby, 2003).

In order to avoid the endogeneity problem, I use a regression model with fixed effects at the municipal level, in which each municipality is compared with itself across time. This prevents underestimates or overestimates of the effect that competition has had on student achievement. Moreover, students in Chile are allowed to attend schools outside to their own municipality. Using the data from the parent questionnaire, I have estimated that around 20 percent of students attend schools outside their municipality in

the capital city (Metropolitan region), and in those municipalities that are located outside of the capital city, around 8 percent of students attend schools outside to their municipality (Simce, 2012). In order to address this situation, I use a dummy variable to identify those students who attend schools outside their municipality, but these data is only available from 2008 to 2012.

An additional problem arises with the way in which competition is defined. Typically, competition is measured using the Herfindahl Index (HI), defined as the sum of the squares of per unit enrollments over total enrollments, bounded between 0 (full competition) and 1 (monopoly) (Borland & Howsen, 1992). Another measure of competition is the private schooling enrollment share, which is the percentage of students enrolled in private-voucher schools in a school district (Dee, 1998; E. A. Hanushek & Rivkin, 2003; C. Hoxby, 1994). In order to make the findings comparable with previous studies in Chile, I use the percentage of students enrolled in 4th grade in private-voucher schools per municipality as a measurement of competition. Moreover, as independent variables I use the student's final grade in his or her class, family income, father's education, mother's education, a dummy variable for years, and a fixed effects at the municipality. Table 3-3 shows the independent variables used in the fixed-effects models.

Table 3-3: Variables Included in the Fixed-Effects Model

	Source	Measured
Dependent variables		
Mathematics achievement test score	SIMCE test	Student's score in math in SIMCE test in 4th grade
Reading achievement test score	SIMCE test	Student's score in reading in SIMCE test in 4th grade
Independent Variables		
Students variables		
Gender	SIMCE test	Student's gender: (1) male (0) female
Transfer	Parent survey	If student enrolls in school outside of their municipality (1) outside and (0) same
Grades	MINEDUC	5 categories: (1) B- (2) B (3) B+(4) A- (5) A
Family Variables		
Family Income	Parent survey	5 categories: from the lowest to the highest income
Father Education	Parent survey	3 categories: (1) primary education, (2) secondary education and (3) college
Mother Education	Parent survey	3 categories: (1) primary education, (2) secondary education and (3) college
School Variables		
Type of school	MINEDUC	(1) Public, (2) Private Voucher (3) Private
Geographical Location	MINEDUC	13 categories: 13 regions and each region is a dummy variable
Urban	MINEDUC	(1) Urban and (0) Rural
SES of students body	MINEDUC	5 categories: From (1) Lowest SES to (5) Highest SES
Municipality		
Percentage of students enroll in private-voucher schools	SIMCE test	Percent of students enroll in private-vouchers school per municipality in 4th grade

I run two different fixed-effects models using a student's score on the math and reading test considering as independent the percentage of students that enroll in private-voucher schools in the Metropolitan and Araucanía regions.

In the next section, I present the main results of the fixed-effects models to estimate the impact of competition on student achievement. The first part presents the

output of the fixed-effects model at the national level and the second section shows the results disaggregated according to the Metropolitan region (capital city) and the Araucanía region.

RESULTS

Table 3-4 shows the output of the fixed-effects model at the municipal level from 2002 (baseline) to 2013, for reading and math. The results point out that the coefficients for competition on reading test, defined as the percentage of 4^{th} grade students enrolled in private-voucher schools per municipality, moves from positive but not statistically significant to negative and statistically significant after controlling for the additional covariantes. A one percent increase in students enrolling in private-voucher school implies a reduction of the students' performance on the reading test by an average of 0.06 points, after controlling for the rest of the independent variables. In the case of the math test, the results show that a one percent increase in competition decreases student achievement on the math test by an average of 0.08 points. Even though the coefficient for competition is significant, its impact on students achievement is very small in terms of the standard deviation of the SIMCE test scores (50 points). The fixed-effects model also shows that male students obtain on average 3.2 points less than female students on the reading test. On the math test, however, male students achieve on average 10.7 points more than female students, after controlling for independent variables. Both coefficients are statistically significant.

Table 3-4: Fixed-Effects Model for 4th Grade Reading and Math Tests at National Level

Variable	reading	reading	math	math
competition	0.01	-0.06***	-0.03***	-0.08***
year				
2005	4.07***	4.14***	0.29*	0.29**
2006	1.81***	1.97***	0.36**	0.76***
2007	3.06***	4.13***	-1.40***	0.04
2008	8.98***	9.86***	-0.14	0.97***
2009	10.46***	11.14***	5.83***	6.92***
2010	19.31***	19.87***	5.79***	6.50***
2011	15.27***	16.02***	11.79***	12.62***
2012	16.21***	17.07***	14.69***	15.56***
2013	12.22***	12.12***	8.90***	8.95***
student's gender		-3.17***		10.75***
student's grades				
B		17.78***		19.01***
B+		39.48***		41.54***
A-		65.32***		69.27***
A		90.47***		97.71***
Father education				
High School		2.97***		2.73***
College		6.34***		5.99***
Mother education				
High School		2.55***		3.13***
College		5.59***		6.01***
Income of family (Monthly)				
US $401-1,000		1.71***		1.83***
US $1,001-2,000		2.26***		2.87***
US $2001-2,300		2.27***		3.81***
More than US $2,300		1.53***		4.66***
School type				
Private-voucher		4.53***		3.60***
Private-paid		1.19***		0.2
Urban		-1.34***		0.83***
School's SES				
Medium low		5.34***		7.16***
Medium		15.69***		18.94***
Medium high		23.81***		28.82***
High		29.15***		36.67***
Cons	251.26***	189.87***	248.67***	171.60***
N	2,287,091	1,829,632	2,287,369	1,835,275

As I expected, as students increase their grades, the mother's education, father's education, the income of their families and school's SES, student scores on the SIMCE test also increase. All of the coefficients are statistically significant at a 0.05 alpha level. Finally, the coefficients by year shows that test scores have been increasing since 2002; there is a clear jump after the implementation of the high-stakes accountability system in 2008. I argue that some of this upward trend can be explained by the implementation of the accountability system

The next section compares regional differences in the impact that competition has had on student SIMCE test scores. As I explained above, I have chosen to compare the Metropolitan region where the capital city (Santiago) is located —the wealthiest and most developed region in Chile — to the Araucanía region, Chile's poorest region and the one with the largest indigenous and rural populations. In this section, I explore the differing impact of competition under these two divergent realities.

Table 3-5 presents the main results of the fixed-effects model at the municipal level across the Metropolitan region. Here the coefficients of the impact of competition on reading and math are positive, which means that for a ten percent increase in student enrollment in private-voucher schools, student scores on the reading test go up by 1.5 points and 1 point on the math test – both coefficients are statistically significant.

As the previous model shows, if students increase their grades, mother's education, father's education and income of their families, student scores on the SIMCE test also increase. All of these impacts are statistically significant. In the case of student gender, the model shows that male students scores, on average, 2.6 points lower than female students on the reading test. In contrast, male students achieve on average almost 12 points more than female students on the math test.

Table 3-5:Fixed-Effects Model of 4th Grade Reading & Math Test in MR

Variable	reading	reading	math	math
competition	0.14***	0.15***	0.06***	0.10***
year				
2005	4.75***	4.46***	1.73***	1.31***
2006	-0.62**	-0.28	1.00***	1.49***
2007	1.26***	2.42***	-0.86***	0.73***
2008	7.20***	8.39***	0.96***	2.22***
2009	7.77***	9.27***	5.47***	7.39***
2010	16.77***	17.97***	5.24***	6.65***
2011	11.58***	13.25***	9.42***	10.96***
2012	12.90***	14.56***	12.35***	13.88***
2013	8.95***	9.48***	7.32***	7.84***
student's gender		-2.63***		11.59***
student's grades				
B		18.47***		19.55***
B+		40.74***		42.58***
A-		66.91***		70.38***
A		90.88***		97.44***
Father education				
High School		3.17***		2.42***
College		5.88***		5.18***
Mother education				
High School		3.11***		3.11***
College		5.93***		5.64***
Income of family (Monthly)				
US $401-1,000		2.36***		2.50***
US $1,001-2,000		3.08***		3.87***
US $2001-2,300		3.35***		5.61***
More than US $2,300		2.81***		6.97***
School type				
Private-voucher		7.01***		6.17***
Private-paid		1.57***		0.39
Urban		-2.45***		-1.31***
School's SES				
Medium low		6.62***		7.79***
Medium		18.58***		20.86***
Medium high		26.61***		30.87***
High		30.52***		36.39***
Cons	247.91***	176.21***	248.13***	161.77***
N	907,250	694,499	907,870	696,631

Furthermore, in relation to the year-specific impacts, we observe an upward trend after the implementation of the SEP program in 2008 (high-stakes accountability), but also a decline in 2013, with the exception of one coefficient for 2006, the rest are statistically significant.

Table 3-6 shows that for the case of the Araucanía region – in contrast to Santiago (in the Metropolitan region) – the coefficients for competition are negative for both reading and math tests. This means that a ten percent increase in competition reduces the student achievement on the reading test by an average of 2.3 points and by 2.7 points on the math test after controlling for the rest of the independent variables. Both coefficients are statistically significant. Similar to the capital city (Metropolitan region), the results for Araucanía region show that for those students with higher grades, as well as higher scores on mother's education, father's education, family's income, and the school's SES, obtain higher scores on SIMCE tests. However, the education of the student's father and mother has a greater effect in Araucanía region than in the Metropolitan region. The impact of family income of the family is also lower in the Metropolitan region, and only statistically significant up to the monthly income category between US $1,001 - $2,000. Finally, the student's gender predicts the same pattern found in the case of Santiago, where male students obtain lower scores on the reading test than female students, while and the opposite is observed in the case of the math test, where male students obtain higher scores.

As I mentioned before, unlike the United States, Chilean students are allowed to enroll in schools outside those in their own municipal school district, and according to the parent questionnaire in 2012 in the capital city, around 20 percent of students in 4^{th} grade enroll in schools outside of the municipality where they live. In the Araucanía region, that percentage decreases to 8 percent.

Table 3-6: Fixed-Effects Model of 4th Grade Reading & Math Test in of Araucanía

Variable	reading	reading	math	math
competition	-0.30***	-0.23***	-0.37***	-0.27***
year				
2005	4.26***	3.86***	-0.86	-0.95
2006	7.54***	5.47***	0.69	-0.19
2007	8.58***	8.10***	-0.51	0.68
2008	14.75***	13.19***	1.65**	1.15*
2009	17.46***	14.04***	10.72***	8.01***
2010	28.75***	24.55***	13.85***	10.03***
2011	26.93***	22.86***	21.20***	17.73***
2012	26.62***	21.85***	23.81***	19.32***
2013	22.05***	16.75***	16.21***	11.53***
student's gender		-2.86***		11.42***
student's grades				
B		17.27***		18.21***
B+		38.99***		40.63***
A-		65.01***		69.16***
A		89.37***		97.93***
Father education				
High School		3.45***		3.62***
College		7.45***		7.85***
Mother education				
High School		2.45***		3.75***
College		5.33***		7.21***
Income of family (Monthly)				
US $401-1,000		1.50***		1.34***
US $1,001-2,000		1.40**		2.63***
US $2001-2,300		-0.31		1.11
More than US $2,300		-1.99*		1
School type				
Private-voucher		-1.27***		-1.62***
Private-paid		2.69		4.48**
Urban		2.85***		7.33***
School's SES				
Medium low		7.83***		10.08***
Medium		16.34***		20.01***
Medium high		21.73***		28.90***
High		27.15***		38.05***
_cons	259.61***	199.19***	257.04***	176.35***
N	126,340	99,677	126,166	99,968

I use this information to create a dummy variable for those students who study outside of their municipality, but the limitation is that the data is only available from 2008 to 2012 and I use the same independent variables in Tables 3-4, 3-5, and 3-6. Table 3-7 presents the most relevant results, showing that even though the period analyzed is shorter than previous models, the results between both models are almost identical (see full table Appendix A-2). In fact, at the national level, the coefficients of competition are very similar; the coefficients for reading and math are both statistically significant and only increase by 0.01 points. In the case of the Metropolitan region, the coefficients are similar to the results in Table 3-5; both are positive and statistically significant.

Table 3-7: Summary of the Main Coefficients of the Fixed-Effects Model that Include Dummy Variable for Those Students Who Attend Schools Outside to Their Municipality

	National		Metropolitan region		Araucanía region	
	Reading	Math	Reading	Math	Reading	Math
Competition	-0.07***	-0.09***	0.14***	0.15***	-0.14*	-0.24***
Enroll outside	0.88***	0.55***	0.89***	0.92***	0.61	-0.59
2008	baseline	baseline	baseline	baseline	baseline	baseline
2009	1.36***	5.95***	0.99***	4.54***	0.91	9.20***
2010	10.06***	5.93***	9.67***	4.31***	11.10***	12.43***
2011	6.23***	11.90***	5.01***	8.51***	9.53***	19.79***
2012	7.32***	14.78***	6.36***	11.41***	8.26***	22.45***
# Cases	885,113	887,987	349,662	350,684	46,908	47,051

For example, the coefficient for competition decreases only 0.01 points in the model for reading and increases 0.05 in the model for math. The coefficients of competition for the Araucanía region are both negative and statistically significant, but the effect of competition on student achievement on the reading test decreases from -0.23 (Table 3-6) to -0.14 points in the current model, and it is only statistically significant at

0.1 alpha level. In the case of math, the coefficient of competition decreases only 0.03 points and is still significant at a 0.01 alpha level.

Finally, at the national level and in the Metropolitan region, the coefficients for those students who enroll in schools outside their municipalities increase their reading scores on average by 0.88 points (national) and 0.92 points (metropolitan region). This is also true for math scores, which increased for those students who attend schools outside to their municipality by 0.55 points (national) and 0.92 points (metropolitan region). All the coefficients are statistically significant. In contrast, the Araucanía region shows that if a student attends a school outside to the municipality where he or she lives, this does not have an effect on his or her performance on either the reading or math tests, since both coefficients are not statistically significant.

DISCUSSION

One of the main assumptions about the implementation of the privatization reform introduced in 1981 was that if parents could exercise their free choice, they would choose the best schools for their children and, therefore, schools would be forced to improve the quality of the education in order to enroll students. However, are Chilean families choosing schools mainly because them provide a good quality of education? To explore this question, the parent questionnaire from 2012 asked parents to choose the three most important variables that they consider in order to choose a school for their children, the questionnaire provided 14 options that parents could select. The questionnaire was a census for 4th grade parents, and 38.4 percent of them reported that the primary variable they take in account in choosing a school is "school proximity" to their home. In addition, 14.3 percent of parents listed that the second most relevant variable they consider in choosing a schools is "school quality and results on SIMCE tests". However,

depending of the socioeconomic status of families, parents found some variables more or less relevant in their decision making process.

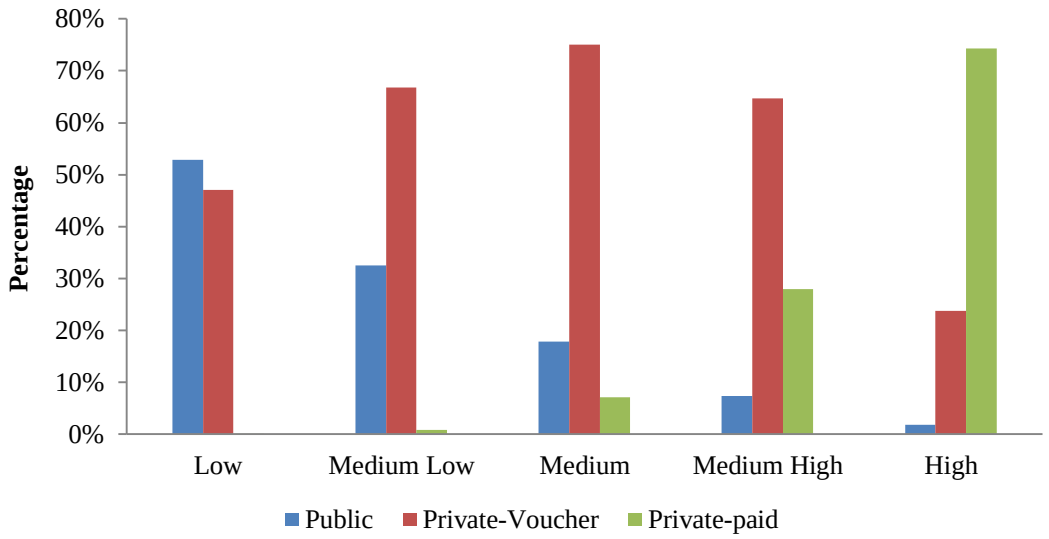
Table 3-8: The Three Main Variables that Parents Consider when Choosing a School, by Socioeconomic Status of Families with Students in 4th Grade, 2012

	1st most important	2nd most important	3rd most important
Family's SES	School proximity	Results of SIMCE tests	School's discipline
Low	63.0%	5.5%	9.6%
Medium Low	55.6%	9.1%	11.9%
Medium	32.5%	19.4%	20.3%
Medium High	20.1%	26.8%	18.9%
High	11.7%	24.0%	8.5%

Table 3-8 shows that as the socioeconomic status of the families increase, "school proximity" become a less relevant variable, while the opposite happens to school quality. As the socioeconomic status of families increase, the "school's quality and results in SIMCE test" becomes more important. And finally, for the middle class the "school's order and discipline" appears to be more relevant.

Moreover, Figure 3-5 shows the relationship between type of school and SES of 4th grade families in in 2013. As is observed, most of the low-income families enrolled their children in public schools, while middle class families enrolled in private-voucher schools and the wealthiest families choose private paid schools.

Figure 3-5: Type of Schools Attended in 4th grade by Socioeconomic Status of Families in 4th grade 2013



Source: Author's calculation using data from 2013 parent questionnaire

In this scenario, the most important aspect of a school for parents is the "school proximity", which is especially relevant for low-income families as well as students that are segregated by the income of their family. This is because unlike public schools, private-voucher schools can select their students through selection tests and they can decide to increase their tuition depending on their own interests and priorities. Therefore, vouchers only cover a certain percentage of the tuition, and parents need to pay extra money to register their children in these schools. Thus, private-voucher schools can improve the average ability level of their student body and create "brain drain" from public schools that would lead to an increase in segregation based on the socioeconomic status of the students (Hsieh & Urquiola, 2002; Levin, 1998; P. J. McEwan, 2000b; Neal, 2011).

As a consequence, the increase of competition does not necessarily mean an increase of quality of education because the possible benefits of competition will be

"cancelled out," due to the increased segregation of the system. In this sense, we have identified an explanation for the negative relationship between competition and student achievement at the national level.

In term of regional differences, I hypothesized that competition has dissimilar effects, dependent upon the conditions where it is implemented. The fixed-effects model for the capital city (Santiago) shows that competition has a modest but positive effect on student achievement. In other words, when the percentage of students enrolled in private-voucher schools in the capital city increases, the reading and math scores of the students also increase. The opposite effect is observed in the Araucanía region, the poorest region of Chile. In this case, when the percentage of students enrolled in private-voucher schools in the municipalities increase, student math and reading test scores decrease. The question that arises is: Why does competition have a different effect depending where it is implemented? To answer, I will focus on one aspect that I think is central to understanding the negative effect of competition in the Araucanía region, and that is the way that rural schools are financed. In particular, I analyze how the way that rural schools are financed has created incentives for private-voucher schools in non-urban areas and how, therefore, the region has observed an increase in competition, but that this has not necessarily meant an improvement in the quality of education.

The Chilean educational finance system is based on the idea that schools need to compete for students and, therefore, it was set up in a way so that public or private-voucher schools received a fixed amount of money for each student (a voucher) according to the student's monthly attendance rate. That means, schools compete for students in order to receive more money and create incentives for schools to encourage students to attend classes. The system was designed more for urban schools, rather than for rural schools, which are usually smaller and also have lower attendance rate due to

geographic conditions and problems associated with transportation. As a consequence, many rural schools were unable to generate enough money to be profitable. With this in mind, in 1998 the Chilean government established a special kind of financial aid for rural schools, called *Piso rural* (rural threshold), which stated that those rural schools with 17 students or fewer would receive a fixed amount of money that will ensure that they have the economic resources to keep the school in operation and that this money would not depend on student attendance rate (Gobierno de Chile, 1998).

This new financial support was beneficial for many under-financed rural schools, but it also create unintended consequences for the proliferation of small private-voucher schools in rural areas. In this sense, even though the competition in the Araucanía region has increased, in a rate similar to that of the Metropolitan region, this is because there are now more private-voucher schools and students enrolled in them. It does not mean the quality of the education has improved. Since the implementation of the rural threshold, owners of private-voucher schools in rural areas prefer to have two schools close to one another, with 15 students each one, rather than open one school with 30 students, because small schools are more profitable.

Chile is divided into 12 regions, with region 1 in the northern part of Chile along the border with Peru and Bolivia, region 12 in the southern region of Chile, with part of its territory in the Chilean Antarctic, and the Metropolitan region (the city of Santiago) located in the central part of the country. The distribution of urban and rural schools is uneven based on the regions; Table 3-9 shows the number of rural primary education schools by region in 2013

Table 3-9: The Percentage of Rural Primary Education Schools by Region in 2013

Region	1	2	3	4	5	MR	6	7	8	9	10	11	12
Urban	64%	87%	69%	42%	78%	93%	47%	37%	49%	31%	28%	54%	74%
Rural	36%	13%	31%	58%	22%	7%	53%	63%	51%	69%	72%	46%	26%

The Araucanía (9) and Lakes (10) regions have the highest percentage of rural primary education schools, at 69 percent and 72 percent in 2013, respectively. One important difference between these two regions is that in 2013, among those rural schools with 17 or fewer students that could receive financial support from rural threshold, only 20 percent of them were private-voucher schools in the Lakes region. Instead, 40 percent of the rural schools with 17 or less students in 2013 in the Araucanía region were private-voucher schools (Mineduc, 2013b). In brief, the Araucanía region has 50 percent more private-voucher schools in rural areas with less than 17 students than its next door neighbor, the Lakes region.

In one of my interviews with an education specialist in the Araucanía region, I asked why private-voucher schools decided to open in rural areas of the Araucanía region? His answer was:

> There is a lot of money, there is a special financial aid called *piso rural* (rural threshold)… school receive a fixed amount of money, which is around CH $1.5 million pesos. If a school has one kid, one teacher, this school will receive CH $1.5 million pesos. Therefore, what happened is that it was better to have 20 schools with 10 students each than one with 200 students, because small schools were more profitable. The owners of private-voucher schools considered this and decided to open a school here, here, here, and there and they can obtain CH $ 6 millions having four schools and this is better business than having one school with 100 students.

Furthermore, rural schools can receive additional funding if they have *jornada escolar completa* (a full day school) or if a school usually runs until 4 PM. Schools could receive an increase in resources, depending on where the schools are located – different regions and municipalities have a different "zone index" which establishes a percent increase in money that schools will receive (Mineduc, 2007b). Finally, even though rural schools have more than 17 students and they cannot receive the "rural threshold," these

schools obtain special financial support depending on how many students attend; the maximum number of students that a school may have to apply for *subvención rural* (rural funding) is 90 students. As a consequence, there are multiple ways that those private-voucher schools can make a profit in rural areas and, as a consequence of that there have been a proliferation of private-voucher schools in the Araucanía region (see Figure A-1), which does not mean an increase in student achievement, as the fixed-effects model for this region shows.

In the Araucanía region, the number of schools in rural areas that closed between the period of 2005 to 2013 was only 32 private-voucher schools versus 127 public schools (Mineduc, 2005, 2013c). Public schools in rural areas face more difficulties in remaining open than private-voucher schools: teachers in the public sector are most costly (Mizala & Romaguera, 2001) and municipalities that are in charge of public schools see them as a problem, given the high cost implied in keeping them open. One extreme example comes from the municipality called Padres Las Casas in the Araucanía region, which closed almost all of its public schools – in fact, 92 percent of urban schools and 79 percent of rural schools were private-voucher in 2013 (Mineduc, 2013c).

In the case of the Metropolitan region, 93 percent of the primary education schools are urban and only 7 percent are rural (Mineduc, 2013c). As a result, competition among schools, in contrast to the Araucanía region, exists among schools in the urban district of the capital city, where parents have more options, where they can exercise their "free choice," and where schools feel pressure to provide a better quality of education in order to attract students. However, despite though all the conditions and advantages that the capital city has, competition has had a very modest contribution to student achievement in terms of the expectations publicly announced by the Pinochet's regimen with the implementation of the education reform in 1981.

CONCLUSIONS

This chapter explored the hypothesis that increased competition, defined as the percentage of students enrolled in 4th grade private-voucher schools per municipality, would result in an improvement in the quality of the education system as a whole. I used longitudinal data from the 4th grade at the student level from 2002 to 2013 to test this hypothesis, and I analyzed the data using a fixed-effects model at the municipal level to avoid endogeneity problems between competition and student achievement.

At the national level, the relationship between competition and student achievement is negative. For example, for a one percent increase in competition, student achievement decreases by 0.06 points on the reading test and 0.08 points on the math test. In fact, these findings are the opposite of what advocates of the implementation of private-vouchers schools have argued. I explain that this effect came about because under the Chilean school choice system, parents identified the most relevant variable in choosing a school to be "school proximity," rather than "quality of education or results in SIMCE test." In fact, the relevancy of "school proximity" is more than 50 percent more relevant than "quality of education" for parents with children in the 4th grade, also private-voucher schools have a mechanism by which they select their student body that not only discriminates against them according to the socioeconomic status of their families, but also according to cognitive abilities. As a consequence, the positive effect of competition and free choice are "cancelled out" by an increase in the socioeconomic segregation of the educational system and freedom that private-voucher schools have to select their student body by income or academic abilities.

Furthermore, this chapter explores the diverse effects of competition on two different regions with dissimilar socioeconomic and demographic conditions. In particular, it compared the Metropolitan region, where the capital city is located, and the

Araucanía region, which is the poorest area of Chile. This study shows that when the percentage of 4th grade students enrolled in private-voucher schools increases in the Metropolitan region, student scores on reading and math tests also increase. This finding provides evidence that in the urban district of the capital city, where families have greater economic advantages than those in the rest of the country, parents can exercise free choice and schools compete to provide a better quality of education. Even though the coefficient of competition is positive and statistically significant, the impact of competition on student achievement is relatively small. For example, a one percent increase in competition implies only a 0.15 point increase on the reading test and a 0.1 point increase on the math test in 4th grade.

In contrast to the urban district of Santiago (Metropolitan region), the coefficient for competition in the Araucanía region, the poorest region with the largest indigenous and rural population, is negative and statistically significant. For example, a one-percent increase in 4th grade students enrolled in private-voucher schools has resulted in a decrease of 0.23 points on the reading test and 0.27 on the math test. I explain this negative relationship between competition and students achievement as due to the negative externalities created by the financial system for rural schools, which creates incentives for private-voucher schools to open in rural areas, but only because they can get more resources from the Chilean government, not because they are competing with others schools.

These findings point to a need for policy makers to take geographical, socioeconomic, and demographic differences into account in the design of public policies. Mainly, because the same public policy could have different effects depending on where it is implemented – if it is with a population that is already better off, the output probably will be positive, but if the implementation takes place in a region where the

population has fared worse than the rest of the country, the expected results are likely going to be limited, or even negative, as the case of the Araucanía region has demonstrated.

References

Ahlin, Å. (2003). Does School Competition Matter? Effects of a Large-Scale School Choice Reform on Student Performance. *Working Paper Series*. Retrieved from http://ideas.repec.org/p/hhs/uunewp/2003_002.html

Altonji, J. G., Huang, C., & Taber, C. R. (2005). *Estimating the Cream Skimming Effect of Private School Vouchers on Public School Students* (Vol. 1). Retrieved from http://www.ssc.wisc.edu/~ctaber/Papers/vouchpap.pdf

Auguste, S., & Valenzuela, J. P. (2004). *Do Student benefit from School Competition? Evidence from Chile*. University of Michigan.

Banco-Central. (2015). Producto Interno por Region. Retrieved June 2, 2015, from http://www.bcentral.cl/estadisticas-economicas/series-indicadores/index_aeg.htm

Belfield, C. R., & Levin, H. M. (2002). The Effects of Competition Between Schools on Educational Outcomes: A Review for the United States. *Review of Educational Research*, *72*(2), 279–341. doi:10.3102/00346543072002279

Bellei, C. (2007). Expansion of private schools and educational improvement in Chile . *Revista Pensamiento Educativo*, *40*(1), 1–21.

Bellei, C. (2012). *Equidad Educativa y Dinámicas de Mercado en Educación*. Santiago, Chile.

Borland, M. V, & Howsen, R. M. (1992). Student academic achievement and the degree of market concentration in education. *Economics of Education Review*, *11*(1), 31–39. Retrieved from http://www.sciencedirect.com/science/article/pii/027277579290019Y

Casen. (2009). *Encuesta de Caracterización Socieconómica Nacional*. Retrieved from http://www.ministeriodesarrollosocial.gob.cl/casen2009/

Casen. (2011a). *Resultados Regionales Encuesta Casen 2011: Región de La Araucanía*. Retrieved from http://observatorio.ministeriodesarrollosocial.gob.cl/layout/doc/casen/IX_Araucania_oct13.pdf

Casen. (2011b). *Resultados Regionales Encuesta Casen 2011: Región de Metropolitana de Santiago*. Retrieved from http://observatorio.ministeriodesarrollosocial.gob.cl/layout/doc/casen/ResultadosRegionalesCasen2011_RegionMetropolitana_50e1bbf8db0b1.pdf

Casen. (2013). *CASEN 2013: Situación de la pobreza en Chile*. Retrieved from http://observatorio.ministeriodesarrollosocial.gob.cl/documentos/Casen2013_Situacion_Pobreza_Chile.pdf

Contreras, D., Bustos, S., & Sepulveda, P. (2010). Emerging Evidence on Vouchers and Faith-Based Providers in Education. In F. Barrera-Osorio, H. Patrinos, & Q. Wodon (Eds.), *Emerging Evidence on Vouchers and Faith-Based Providers in Education* (pp. 1349–1368). The World Bank.

Corvalan, J. (CIDE), Elacqua, G. (UDP), & Salazar, F. (UDP). (2009). *El sector particular subvencionado en Chile. Tipologización y perspectivas frente a las nuevas regulaciones*. Retrieved from http://biblioteca.uahurtado.cl/ujah/reduc/pdf/pdf/mfn341.pdf

Dee, T. S. (1998). Competition and the quality of public schools. *Economics of Education Review*, *17*(4), 419–427. doi:10.1016/S0272-7757(97)00040-X

Elacqua, G. (2006). *Enrollment practices in response to vouchers: Evidence from Chile* (No. 125). New York City. Retrieved from http://www.ncspe.org

Flores, C. (2008). *Residential Segregation and the Geography of Opportunities: A Spatial Analysis of Heterogeneity and Spillovers in Education*.

Friedman, M. (1962). *Capitalism and Freedom*. Chicago: University of Chicago Press.

Gallego, F. A. (2002). *Competencia y Resultados Educativos:Teoria y Evidencia Para Chile* (No. 150). Santiago, Chile.

Gallego, F. A. (2004). *School Choice , Incentives , and Academic Outcomes: Evidence for Chile* (No. 39). Retrieved from http://ideas.repec.org/p/ecm/latm04/39.html

Gobierno de Chile. Decreto con Fuerza de Ley 2 (1998). Retrieved from http://www.leychile.cl/Navegar?idNorma=127911

Gobierno de Chile. (2011). Plan Araucania 7. Retrieved December 10, 204AD, from http://araucania7.cl/el-plan/

González, P., Mizala, A., & Romaguera, P. (2004). *Vouchers, Inequalities and the Chilean Experience*. Santiago, Chile. Retrieved from http://www.ncspe.org/publications_files/Voucher_ChileOP94.pdf

Hanushek, E. A., & Rivkin, S. G. (2003). The Economics of School Choice. In C. M. Hoxby (Ed.), (Vol. I).

Hoxby, C. (1994). *Do Private Schools Provide Competition For Public Schools?* (No. 4978). Retrieved from http://www.nber.org/papers/w4978

Hoxby, C. (2000). Does Competition Among Public Schools Benefit Students and Taxpayers? *The American Economic Review*, *90*(5), 1209–1238. doi:10.1257/aer.97.5.2026

Hoxby, C. (2003). School choice and school competition: Evidence from the United States. *Swedish Economic Policy Review*, *10*, 9–65.

Hsieh, C., & Urquiola, M. (2002a). *When schools compete , how do they compete ?* (No. 43). New York City.

Hsieh, C., & Urquiola, M. (2002b). *When Schools Compete, How Do They Compete? An Assessment Of Chile's Nationwide School Voucher Program*. *NBER Working Paper Series* (Vol. 10008). doi:10.1016/j.jpubeco.2005.11.002

Ladd, H. F. (2002). School Vouchers: A Critical View. *Journal of Economic Perspectives*, *16*(4), 3–24.

Levin, H. (1998). Educational vouchers_ Effectiveness, choice, and costs.pdf. *Journal of Policy Analysis and Management*, *17*(3), 373–392.

Levin, H. (2002). A Comprehensive Framework for Evaluating Educational Vouchers. *Educational Evaluation and Policy Analysis*, *24*(3), 159–174.

McEwan, P., & Carnoy, M. (1999). *The Impact of Competition on Public School Quality: Longitudinal Evidence from Chile's voucher System* (No. Unpublished Manuscript).

McEwan, P. J. (2000). *The Potential Impact of Large-Scale Voucher Programs*. *Review of Educational Research* (Vol. 70).

Mineduc. (2005). Matricula por Estudiante. Retrieved January 2, 2015, from http://centroestudios.mineduc.cl/index.php?t=96&i=2&cc=2036&tm=2

Mineduc. (2007). *Guía de Subvenciones Educacionales*. Santiago. Retrieved from http://w7app.mineduc.cl/LiquidacionSubv/Principal/guia.pdf

Mineduc. (2013). Matricula por Estudiante. Retrieved January 2, 2015, from http://centroestudios.mineduc.cl/index.php?t=96&i=2&cc=2036&tm=2

Mizala, A., & Romaguera, P. (2000). School Performance and Choice The Chilean Experience. *The Journal of Human Resources*, *35*(2), 392–417.

Mizala, A., & Romaguera, P. (2001). *Regulación, incentivos y remuneraciones de los profesores en Chile*. CEA. Retrieved from http://books.google.com/books/about/Regulacion_incentivos_y_remuneraciones.html?id=Jb6RYgEACAAJ&pgis=1

Neal, D. (2011). How Vouchers Could Change the Market for Education. *The Journal of Economic Perpectives*, *16*(4), 25–44.

Nechyba, T. J. (2003). Introducing School Choice into Multidistrict Public School Systems. In C. M. Hoxby (Ed.), *The Economic of School Choice* (pp. 145–192). University of Chicago Press.

Paredes, R., & Pinto, J. I. (2009). Is this the end of Public Education in Chile? *Estudios de Economia*, *36*, 47–66.

Torche, F. (2005). Privatization Reform and Inequality of Educational Opportunity : *Sociology of Education*, *78*(Mmi), 316–343.

Urquiola, M. S. (2000). Demand Matters: School District Concentration, Composition, and Educational Expenditure. *The World Bank*. doi:10.2139/ssrn.244099

Valenzuela, J. P., Bellei, C., & De Los Rios, D. (2008). *Evolución de la Segregación Socioeconómica de los Estudiantes Chilenos y su Relación con el Financiamiento Compartido*.

Chapter 4

The Use of Academic Strategies by Teachers under the High-Stakes Accountability System in Chile

Diverse studies have found a positive effect of the implementation of high-stakes accountability system implemented in Chile in 2008, the so-called *Subvención Escolar Preferencial* (Preferential Voucher Program, SEP), and a resulting improvement of the quality of education (Correa et al., 2012; Mineduc, 2012; Mizala & Torche, 2013; Perticara et al., 2013; Villarroel, 2012). In others words, schools that voluntarily enter the Preferential Voucher Program (SEP) have shown an increase in students' performance in the *Sistema Nacional de Evaluación de Calidad de la Educación* (Educational Quality Measurement System, SIMCE), but it is not possible to know from these studies how or what schools did in order to improve their performance. As Koretz (2008) points out, teachers respond differently to the high-stakes accountability system; some of them develop strategies that are desirable and expected from the implementation of an accountability system, such as so-called academic strategies. For example, it is expected that teachers spend more time with students, find new learning methods, or give students more homework or assignments, and as a results of that students improve their performance on the standardized test. However, there are other strategies (non-academics strategies), such as teaching to the test or excluding low-performing students from the test-taking pool, which also serve to improve students' scores on the standardized tests, but these non-academic strategies are not the type of actions that the accountability system is expecting from teachers.

This chapter analyzes the use of academic strategies following the implementation of the high-stakes accountability system in Chile; academic strategies are those that that teachers or schools are expected to develop as a way to improve the students' learning

and, as a result, to observe better performance by students on standardized tests. These strategies are (a) teachers work more effectively: for example, they find better methods of teaching, (b) teachers teach more: teachers spend more time overall in school, and (c) teachers work harder: for example, they give more homework or harder assignments to their students (Koretz, 2008a). These three strategies produce higher levels of achievement in substantive areas of the test and an unambiguously meaningful gain in scores (Koretz & McCaffrey, 2001).

From the current literature is impossible to quantify how extensive or significant the use of academic strategies are under the SEP accountability system. It is not possible to know what type of municipalities, schools, labor conditions, or which characteristics of principals help to develop these types of strategies to improve students' learning and, therefore, their performance on standardized tests. In this sense, this study is a contribution to the literature, but it also offers fundamental information to policy-makers in Chile about the design and implementation of the nationwide high-stakes accountability system that the Chilean government would put into place in 2015.

This chapter explores the following research questions: after the implementation of the high-stakes accountability in Chile in 2008, have teachers increased the use of academic strategies as a way to improve students' performance on standardized tests? And what are the difficulties and constraints that the Chilean educational system imposes on teachers and school administrators in order to increase the use of academic strategies under the high-stakes accountability system?

In the next sections, I introduce an overview of how incentives are understood in economics and psychology, and how these theoretical notions were introduced and used to implement the high-stakes accountability system in Chile in 2008.

THEORY OF INCENTIVES: PERSPECTIVES FROM ECONOMICS AND PSYCHOLOGY APPROACHES

In past decades, economists and psychologists have studied the complex relationship between incentives and human behavior (Hout & Stuart, 2011). The field of economics is based on incentives and its strength comes from being able to predict how human beings modify their behavior in response to changing incentives, but economics are also more inclined to constrain their attention to narrow and empirically-verifiable aspects of human motivation (Fehr & Falk, 2002). While psychology tries to explain the motivation behind people's actions, economics tries to explain human motivation and how individuals perceive incentives such as rewards or sanctions could modify their behavior (Benware & Deci, 1984).

How to design institutions in order to provide good incentives for individuals is a fundamental question in economics. Neoclassical economics argues that individuals behave rationally in the market, trying to maximize their utilities and minimize their cost. In this sense, the pressure of competitive markets solves the problem of incentives for cost minimization because in a perfect competitive market, consumers faced with exogenous prices have the proper incentives for maximizing their utilities levels. As a consequence, an understanding about how prices are formed in competitive markets is achieved without worrying about incentives. The problem is, when firms are analyzed as a black box, it is difficult to understand how and why some owners of firms succeed in aligning their workers, supervisors and managers with profit maximization. When economists become more interested in understanding how firms work, incentives arise as a central aspect of this analysis (Laffont & Martimort, 2001).

In order to develop a theory of incentives, it is necessary to address the problem of delegating a task to an agent with private information, where one party (the principal)

tries to influence the behavior of another (the agent) by using incentives. In this case, there are two types of private information: The first one, called a "moral hazard," is a situation in which the agent's actions affect the principal's payoff, even though the action is not directly observable to the principal. The second one is called "adverse selection," which corresponds to a case of information asymmetry where the agent has private information that is ignored by the principal (Dixit, 2002). One of the critiques of this approach is related with the notion of human behavior; the principal-agent relationship has been developed based on the notion that agents are fully rational maximizers. However, in some cases, agents could fail to optimize their outcome and display some irrational behavior (Laffont & Martimort, 2001).

Does the optimal structure of incentives in the public sector differ from those in the private sector? One of the main characteristics of the public sector is that agents work for many principals, and to create incentives in this context is more complex. Additionally, the principals are usually interested in some but not all dimensions, and those interests are not necessarily aligned as well, as most of those employed in the public sector perform several tasks (Burgess & Ratto, 2003; Dixit, 2002). In this context, negative externalities appear, as the different principals expect different outcomes from the agents, and each principal only rewards or sanctions those outcomes in which he or she is interested. Therefore, the agents will be more willing to put more effort and time into satisfying those outcomes that have a higher reward for them and, therefore, will tend to disregard the rest of the tasks associated with their job (Dixit, 1997).

In the case of public education, the system presents multiple goals. For example, schools need to teach basic skills of mathematics, sciences, arts, reading, reasoning and calculation; foster physical and emotional growth of children; teach students ideal of citizenship and responsibility; provide an environment free of drugs and violence.

Furthermore, public education has several principals such as parents, teacher unions, taxpayers, and society as a whole, and teachers and school administrators (agents) need to work to satisfy them. Another feature of the system is that teachers have long careers, which means that teachers can be more focused on the long-term goals of their career rather than incentives based on short-term performance (Dixit, 2002). In summary, these characteristics of public education, which is multi-task, multi-principal, and multi-period, creates almost a monopoly organization with difficult observable outcomes and it is difficult to apply a single solutions as the theory of incentives would propose (Burgess & Ratto, 2003).

An additional theoretical source that explores the relationship between incentives and behavior comes from the field of psychology. This approach reveals and underlines the importance of the desire to reciprocate or the desire to avoid social disproval in human behavior. In disregarding these motives, economist could fail to fully understand the levels and changes in human actions (Fehr & Falk, 2002). From this approach, to understand the effect of incentives is central in considering the interpretation of the rewards or sanctions that the recipients are attaching to the incentives. Particularly, the recipient's perception of the incentives in relation to their own feeling of self-determination (Deci et al., 1999). A person's interpretation of the causes of behavior are hypothesized to have an important impact on future motivation and performance (Cameron & Pierce, 1994).

According to cognitive evaluation theory, all incentives and constraints have two functional aspects: a controlling aspect and an informational aspect (Deci, Schwartz, Sheinman, & Ryan, 1981). The controlling aspect is referred to as how the locus of causality is perceived. For example, when individuals are intrinsically motivated to do an activity, they perceive that locus of causality as internal and therefore people feel self-

determined. In contrast, if people are extrinsically motivated, they interpret that locus of causality as external and feel less self-determining. The informational aspect refers to the idea that intrinsic motivation can be influenced by the way people perceive their own competence. If people perceive that they have high levels of competence, their intrinsic motivation will increase under situations that require them to meet certain goals or standards. The opposite happens if people's interpretations or feelings of competence are diminished: their intrinsic motivation will also will decrease (Deci, Nezlek, & Sheinman, 1981).

Cognitive evaluation theory understands intrinsic motivation as the psychological needs for autonomy and competence, and therefore, the influence of an event such as a reward depend on how the individual perceives it, and how they perceive their own self-determination and competence. Intrinsically, motivation includes behaviors associated with exploration, spontaneity, enjoyment, and excitement. In contrast, extrinsic motivation involves performing activities with the goal of receiving a reward, avoiding sanction, or gaining approval. Those actions that are extrinsically motivated would be less spontaneous and because this type of behavior is focused on reaching an objective, the behavior is instrumental (Deci, Ryan, & Williams, 1996). According this theory predicts that when an interpersonal style of administering performance-contingent rewards is relatively pressuring, the rewards are most likely perceived as more controlling, and therefore leading to a reduction of intrinsic motivation. On the other hand, when the interpersonal style is not controlling, the rewards tend to be experienced as more informal, and could lead to an increase of intrinsic motivation (Deci et al., 1999; Deci, Schwartz, et al., 1981).

I argue that under the high-stakes accountability system implemented in Chile (2008), teachers will respond using academic and non-academic strategies based on their

perception of self-determination and competence. In others words, if teachers perceive the goals defined by the accountability system as beyond their capabilities, as cognitive evaluation theory predicts, they will develop instrumental behavior (extrinsic motivation) to achieve the goals established by the accountability system. I hypothesize that they will be most likely to develop non-academic strategies to improve students' performance in standardized tests. On the other hand, if teachers perceive the objective defined by the accountability system as something they can achieve (intrinsic motivation), they will most likely develop academic strategies to improve the students learning and, therefore, students' performance on standardized tests. Moreover, I argue that the way teachers perceive their competence is determined by their skills and the school's previous performance on the SIMCE tests. That is, if a teacher possesses a good set of skills and the school has had good performance on the SIMCE test (lower pressure for results), most likely this teacher will be intrinsically motivated and prone to use academic strategies. In contrast, if a teacher has low skills and the school has obtained low scores on SIMCE test (higher pressure for results), most likely teachers will be extrinsically motivated to obtain results and inclined to use non-academic strategies.

The High-Stakes Accountability System in Chile

The use of standardized tests is understood as a necessary aspect of improving the quality of education. With this in mind, in recent decades in the United States, the efforts of federal and state governments have been focused on implementing accountability systems in education (Chapman et al., 2000; Cullen & Reback, 2006; Figlio & Getzler, 2002; Hamilton et al., 2002; E. Hanushek & Raymond, 2005; Jacob, 2007; Kim, 2004; Neal & Schanzenbach, 2010; Stecher & Barron, 2001). It is assumed that the implementation of the standardized test as a mechanism of rewards and sanction would

be a source of motivation for teachers and schools administrators to improve their work and maximize the students' learning. From this prospective, an increase in expectations and learning standards should promote better opportunities for students and reduce the academic achievement gap among students (Haertel & Herman, 2005).

While the U.S. educational system has a lower degree of competition among schools (measured as the percentage of students enrolled in voucher and charter schools in school districts), some states have implemented a high-stakes test-based accountability system. In contrast, Chile had a high level of competition among schools paired with the use of low-stakes testing up to 2007. More than half of students attend private-voucher schools, but standardized tests and their results are mainly employed to help inform parents about school quality, and they have no direct consequences for teachers or school administrators (Mizala & Urquiola, 2007). However, this situation changed in 2008, when the Chilean government implemented the Preferential Voucher Program, also known as SEP. The original voucher program, implemented in 1981 under the military regime, established a flat voucher for all students, disregarding the fact that students from low-income families would require additional resources to achieve similar results that their peers from high-income families (Mizala, 2007). To solve this problem, the Chilean government created the Preferential Voucher Program, which gives additional resources to schools by increasing the number vouchers for students from low-income families, but also established a high-stakes accountability system in which schools must achieve certain goals as defined by standardized tests (Romaguera & Gallegos, 2010). Even though SEP is voluntary, in 2011, three years after its implementation, 99 percent of the public schools and 73 percent of the private-voucher schools participated in SEP program and the Chilean government had invested around $1,300 million dollars on this program (Mineduc, 2012).

Various studies have evaluated the impact of the Preferential Voucher Program (SEP) on students' achievement in SIMCE tests. Scholars have used different techniques, such as propensity matching scores or difference in difference (Correa et al., 2012; Mineduc, 2012; Villarroel, 2012). The problem of using these techniques is that schools were not chosen to participate in the program; instead, schools voluntarily entered the program (Perticara et al., 2013). As a consequence, it could be the case that those schools with better students and parents who are more motivated would be more likely to participate in the SEP program and, therefore, if SEP schools (treatment group) are compared with those that do not participate in the program (control group), researchers could find a positive effect that would erroneously be attributed to the Preferential Voucher Program (SEP). On the other hand, if schools with more disadvantaged students and less motivated parents were most likely to join the program, researchers who compare SEP schools versus non-SEP schools would conclude that the high-stake accountability system has a negative impact on the students' performance on the standardized tests, even though the differences are related with the composition of the student body (Mizala & Torche, 2013). Because of the problems described above, another other group of researchers has chosen to use a fixed effects to evaluate the effect of the high-stakes accountability system over the test-scores model (Mizala & Torche, 2013; Perticara et al., 2013). These studies compare schools against themselves over time, including observations before and after the implementation of the SEP program.

Even though studies have used different approaches to explore the effect of the Preferential Voucher Program, they have obtained similar conclusions. Those schools that have participated in the SEP program have increased student performance on standardized tests (Correa et al., 2012; Mineduc, 2012; Mizala & Torche, 2013; Perticara et al., 2013; Villarroel, 2012). However, it is unclear how or what schools did in order to

improve their test scores, given that I have explained previously how teachers can use academic and non-academic strategies in order to improve student's scores on standardized tests. Academic strategies include finding new learning methodologies to teach, spending more time with students, or giving more homework or assignments to students. On the contrary, non-academic strategies include coaching or teaching to the test, selecting students or excluding low-achievement students from the test-taking pool (Koretz, 2008b).

This chapter explores whether or not teachers use mostly academic strategies to improve student performance on standardized test, as the theory assumes. Previous researches have analyzed this question. For example, in 1999 Eyzaguirre and Fontaine pointed out that the accountability system could have positive and also negative consequences (Eyzaguirre & Fontaine, 1999). Moreover, Roman argued that one of the challenges was to control the negatives effects of the Chilean accountability system (low-stakes), such as the selection and exclusion of the low-achievement students (Roman, 1999). In 2003, the Minister of Education created a commission to study the impacts of the low-stakes accountability system implemented in Chile in the 1990s. The commission found evidence of reallocation, or that schools tended to concentrate their efforts on those grades that will take the standardized tests, while teachers became more focused on teaching math and reading, disregarding the rest of the subjects. Moreover, the report showed that schools put their best teachers in charge of the grades that would take the standardized tests and that they would align the curriculum with the topics that the standardized test would be covering (Mineduc, 2003).

After the implementation of the high-stakes accountability system in 2008, some studies have used qualitative methods to discover that schools are matching curriculum more closely to the material covered on the test (alignment) while also teaching to the test

(Falabella, 2014; Waissbluth, 2013). However, it is unclear from these studies how extensive and significant these practices are or whether it could be the case that academic strategies have also increased following the implementation of the high-stakes accountability system. This study tests the hypothesis that advocates of accountability favor: that under the pressure and objectives set up by the high-stakes accountability system, teachers will increase the use of academics strategies as a way to improve student performance on standardized tests.

METHODS AND DATA

I use census data from a questionnaire given to 4th and 8th grade teachers. This data is restricted and it is necessary to apply to and be authorized by, the Ministry of Education in order to have access to it. The teacher of each class answered the questionnaire, which inquires about his or her professional training and the contents that were taught during the school year, among others. The majority of the questionnaires were answered by one teacher who taught the three classes tested by SIMCE: math, reading, and sciences (Mineduc, 2008).

As table 4-1 shows, there is information available from the 4th grade over 10 periods from 2002 and 2005-2013 and from 8th grade over three periods in 2007, 2011 and 2013. Each teacher's questionnaire has a unique ID, which I have used to merge the information about a teacher's schools. When both datasets are merged, I can see whether or not a particular teacher is working in a SEP or non-SEP school, what class he or she is teaching, the socioeconomic status of students in his or her school, the geographical location of his or her school, the type of school (public, private-voucher or private-paid) in which each he or she teaches.

Table 4-1: Total Numbers of 4th and 8th Grade Teachers who Answered the Questionnaires from 2002 to 2013

	2002	2005	2006	2007	2008	2009	2010	2011	2012	2013
4th grade	9,061	10,560	10,526	10,732	10,607	9,890	10,600	9,981	10,190	10,064
8th grade				18,055				16,682		14,982

Furthermore, I analyzed 23 semi-structured interviews with academics, policy-makers and teachers in Chile. These interviews were conducted in two periods: from September to December of 2013 and in August, 2014. The first groups of interviews conducted in 2013 was focused on describing the main problems that the Chilean educational system has faced since the implementation of the private-voucher reform under the dictatorship of Pinochet in 1981, and what challenges it faced after the implementation of the high-stakes accountability system in 2008. The second round of interviews conducted on August 2014 focused on exploring the use of academic and non-academic strategies for SEP schools and the difficulties they face in utilizing academic strategies to improve students' learning. I use descriptive statistics such as frequency diagrams and contingency tables to analyze the teacher questionnaires, and in the case of the semi-structured interviews, I use content analysis.

RESULTS

This section explores teachers' responses to high-stakes accountability system, based on the typology created by Daniel Koretz (2008). In particular, I analyze whether or not academic strategies are taking place after the implementation of the SEP program in Chile. I use the teacher questionnaires to explore three expected actions that teachers could develop under the accountability system: (1) working more effectively, (2) teaching more, and (3) working harder. These three types of responses can produce higher scores

on standardized tests and ensure the inference that students have improved their understanding or learning of the intended domain (Koretz & McCaffrey, 2001).

Finally, I analyze 23 semi-structured interviews conducted in Chile from October to December 2013 and from August 2014 with the goal of understanding the constraints in the Chilean system that make it difficult for schools to choose academic strategies as a way to improve their student performance on standardized tests.

Working More Effectively

This type of teacher responds to the high-stakes accountability system by finding better methods of teaching or by adopting a better curriculum (Koretz & McCaffrey, 2001). I use the 8th grade teacher questionnaire to see if teachers who work in SEP schools have increased the diversity of evaluations, feedback, or methods of teaching their students since the implementation of the SEP program. As I mentioned before, it is important to clarify that SEP and non-SEP schools are not comparable, because the SEP program is not a mandatory program; because schools choose whether or not to participate in this program and are not randomly assigned to be either an SEP or non-SEP school, the groups are not comparable. However, I use SEP and non-SEP schools to observe whether or not they are following the global national trend, or if SEP schools are developing their own trajectory after the implementation of the high-stakes accountability system in 2008.

In order to explore the type of evaluation that schools are using after the implementation of the high-stakes accountability system, Tables 4-2 and 4-3 show the proportion of written and multiple choices tests given in 8th grade prior to and following the implementation of the SEP program. I use these two types of evaluations because I hypothesize that SEP schools will reduce the use of multiple-choice tests. Otherwise, it is

possible to suspect that there could be evidence of coaching, as students could be receiving training as to how to answer questions in a format similar to that of the SIMCE tests. Moreover, I expect to observe some increase in the use of written tests in SEP schools, because if SEP schools are seeking to improve the learning process and not just prepare their students to take standardized tests, then written tests are a better evaluation of whether or not students comprehend tests questions, the use of abstract thought, and analysis in their answers. For example, using tests with multiple choice items, students who don't know the answer know that they can find the correct response they by using information in neighboring items (Pyrczak, 1972).

I use the years before of the SEP program implementation (2008) as a baseline and I compare the four years after the implementation of the high-stakes accountability system. Table 4-2 shows answers from 8th grade teachers that use multiple-choice tests. The results demonstrate that SEP schools have decreased the percentage of teachers who are evaluating their students using multiple-choices tests, which is what I hypothesized previously, but also that non-SEP schools reduced the use of multiple-choice testing after the implementation of the high-stakes accountability system. In fact, the percentage of teachers that report to "always" use multiple choice tests decreased on average around 7.7 percent between 2007 and 2011 among SEP schools, but non-SEP schools increased by 0.9 percent in the same period. Moreover, the percentages of teachers that say "never" use multiple-choice tests in reading class decreased by 2.2 percent for SEP schools and 7.5 percent for non-SEP schools.

Table 4-2: In Reference to this Class (8th Grade), How often do you use Tests with Multiple-Choices or True & False Items in Reading?

	Always		Never	
	SEP	NON SEP	SEP	NON SEP
2007	43.1%	40.5%	3.3%	9.7%
2011	35.4%	39.6%	1.1%	2.2%
Difference	-7.7%	-0.9%	-2.2%	-7.5%

In the case of tests where students need to write their answers, Table 4-3 shows that SEP and non-SEP schools follow similar trends. In both cases, the percentage of teachers that "always" use written tests decreased 4.5 percent in SEP schools and 5.2 percent for non-SEP schools after the implementation of the accountability system.

Table 4-3: In Reference to this Class (8th Grade), How often do you use Tests where Students Need to Write their Answers or Solve Problems in Reading?

	Always		Never	
	SEP	NON SEP	SEP	NON SEP
2007	43.1%	50.9%	1.3%	1.3%
2011	38.6%	45.7%	0.3%	0.2%
Difference	-4.5%	-5.2%	-1.1%	-1.1%

In summary, the type of evaluations that schools are using four years after the implementation of the high-stakes accountability system show that SEP schools have reduced multiple-choice tests as a way to measure students' learning, which is the outcome I had hypothesized. But the percentage of teachers that reported to "always" evaluate their students using written tests also decreased after the implementation of SEP program, what is the opposite of what I was expecting from an SEP school. Moreover, non-SEP schools have followed the same pattern as SEP schools after the implementation of high-stakes accountability system.

Another element that I explore in this section is the idea that whether or not SEP schools are intensifying the feedback that teachers are giving to students, as a way to improve and accelerate students' learning (Nicol & Macfarlane-Dick, 2006). There is no data available before 2008 and I therefore used 4th grade data from the period between 2009 and 2013 to explore whether or not teachers from SEP schools have increase the frequency of feedback they give to their students. Table 4-4 shows that SEP schools have increased the percentage of feedback given to students on study guides, with homework, and that their feedback on tests decreased

Table 4-4: Percentage of Teachers in 4th Grade from SEP & Non-SEP Schools that "Always" or "Never" Give Feedback to their Students in 2009 and 2013.

		Always		**Never**	
		SEP	NON SEP	SEP	NON SEP
Study guides	2009	58.9%	67.4%	0.1%	0.1%
	2013	71.9%	72.7%	0.2%	0.2%
	Difference	13.0%	5.4%	0.04%	0.2%
Homework	2009	53.6%	57.3%	1.0%	0.8%
	2013	66.5%	59.2%	0.4%	0.8%
	Difference	12.9%	2.0%	-0.65%	0.0%
Tests	2009	59.6%	67.4%	0.2%	0.1%
	2013	47.0%	52.6%	1.6%	3.4%
	Difference	-12.6%	-14.8%	1.43%	3.3%

For example, in 4th grade classrooms where teachers who work in SEP schools reported that "always" give feedback to students based on study guides represented 59 percent in 2009 and 72 percent in 2013, showing an increase of 13 percent. In contrast, teachers in non-SEP schools increased by only 6 percent from 67 percent in 2009 to 73 percent in 2013. Moreover, the percentage of teachers from SEP schools that "always" give feedback to student about their homework was 54 percent in 2009 and 66 percent in 2013, which represent an increase of 12 percent. On the other hand, the percentage

increase of among non-SEP schools was only 2 percent, from 57 percent in 2009 and 59 percent in 2013. Finally, both SEP and non-SEP schools have seen a reduction in the percentage of teachers that "always" give feedback to their students on tests, with SEP schools shifting from 60 percent in 2009 to 47 percent in 2013 while non-SEP schools moved from 67 percent in 2009 to 53 percent in 2013.

In summary, even though SEP and non-SEP schools follow similar trends in relation to the feedback that students receive, SEP schools show a higher increase in the percentage of teacher that "always" give feedback to their students in study guides and homework. In fact, in 2013 SEP schools show a higher percentage of teachers that give feedback in homework to their students: 66 percent of SEP schools and 59 percent of non-SEP schools. They had similar percentages in terms of teachers that "always" gave feedback about study guides to their students: 72 percent for both types of schools.

In relation of the use of different methods of teaching, if SEP schools are more focused on students learning rather than preparing them to take standardized tests, it could be expected that one would observe an increase in the percentage of teachers that report the use of more personalized work with students, such as methods of teaching focused on individual and group work. Table 4-5 compares the percentage of teachers that "always" used different methods of teaching before the implementation of SEP program (2007) to those six years after the SEP reform (2013). The results show that SEP and non-SEP schools follow similar trends, with the percentage of teachers that report to "always" use the eight methods of teaching present in Table 4-4 declining after the implementation of SEP program in 2008. The most important declines can be attributed to methods that involve group or individual work with students. For example, teachers from SEP schools that "always" do group work was 21 percent in 2007 and 7 percent in 2013, which is a decline of 14 percentage points. On the other hand, non-SEP schools

show a 9 percentage points reduction in group work over the same period, from 15 percent in 2007 to 6 percent in 2013.

Table 4-5: The Percentage of Teachers from SEP & Non-SEP Schools who Report that they "Always" Use Different Methods of Teaching in 8th Grade, in 2007 & 2011

		Always	
Methodology		SEP	NON SEP
Group activities	2007	20.6%	14.6%
	2013	7.0%	6.2%
	Difference	-13.6%	-8.4%
Individual with students	2007	37.4%	40.1%
	2013	17.4%	18.1%
	Difference	-20.0%	-21.9%
Lecture of teacher	2007	23.6%	24.0%
	2013	20.3%	21.2%
	Difference	-3.3%	-2.8%
Questions & Answers	2007	28.8%	29.7%
	2013	26.9%	25.0%
	Difference	-1.9%	-4.7%
Student presentations	2007	13.3%	6.9%
	2013	7.5%	5.0%
	Difference	-5.8%	-1.9%
Debate among students	2007	10.6%	5.7%
	2013	4.7%	4.6%
	Difference	-6.0%	-1.1%
External activities (parks, museums, etc.)	2007	5.4%	3.7%
	2013	3.5%	2.7%
	Difference	-1.9%	-1.1%
Students need to present a final project	2007	1.8%	1.0%
	2013	1.2%	0.6%
	Difference	-0.6%	-0.4%

Furthermore, teachers from SEP and Non SEP schools report a decline of almost 16 percent in the use of individual work with students, when comparing the years 2007 and 2013. For example, 37 percent of teachers in SEP schools indicated that they "always" did individual work with students in the classroom in 2007 versus 17 percent in

2013. On the other hand, 40 percent of teachers in non-SEP schools reported that they "always" did individual work with students in 2007 and 18 percent did in 2013. Moreover, Table 4-5 shows that the use of teaching methods that include lecture, activities focus on question and answer, exposition, debate, external activities and final project has decreased after the implementation.

Table 4-5 has shown that teachers are not using clear methods of teaching in their classes, based on the finding that teachers who selected "always" decreased from 2007 to 2013, and those selecting "sometimes" increased after the implementation of the high-stakes accountability system. This could be positive, as it could show that teachers are more flexible in using different methods to teach their students, but also it could mean that there is no clear guidelines about what methods of teaching at the SEP schools should be used in order to ensure better and deeper learning among their students. A study conducted by the Catholic University of Chile (PUC) in 2014 showed that principals determine the rotation of teachers, which means that teachers do not stay for a long period of time in the same school, a factor that could explain the use of different methods of teaching (Manzi, Bogolasky, Grau, Guitiérrez, & Volante, 2014).

In an interview with a teacher who works in a private-voucher school, he made the following comments with regards to the labor market of teachers:

> I've been in a school for up to two years, but in the second year they don't rehire you because in the second year, by law, they would have to give you a permanent contract. And the permanent contract doesn't give the owner liberty to fire you, because they would have to compensate you. This labor instability is for all teacher who work in private-voucher schools.

Since the educational reform implemented under the dictatorship, rules and norms that govern a teacher in the public and private sector are different. Teachers who work in private-voucher schools lose their condition as a public employee and are instead

regulated under the private labor code (código del trabajo) that governs the private sector in Chile. In this system, there are no considerations for how the salaries or a teacher's career should be regulated in the private sector (Mizala & Romaguera, 2001). In this sense, private-voucher schools behave much as private company would. They are allowed to contract teachers using a temporary contract of a maximum of two years, after which time they must shift to a permanent contract that would require them to pay teachers' vacations and include other benefits. As a consequence, private-voucher schools try to avoid these extra costs by firing teachers after two years and hiring new ones. This element could partially explain the relative fluctuations in methods of teaching observed in Table 4-5, while also it is important to note that private-voucher schools educate more than half of Chilean students. For example, in 2013 private-voucher schools enrolled 55 percent of students in primary and secondary education (Mineduc, 2013b)

Teaching More: Spending more time overall

A second type of response that teachers can develop under a high-stakes accountability system, according to Koretz, is spending more overall time working with their students. I used the teacher questionnaires from the years 2006, 2009, and 2010 to explore whether or not teachers form SEP schools spend more time preparing their classes than they used to before the implementation of the SEP program. Table 4-6 shows the average time (hours) spent by teachers preparing their classes before (2006) and after the implementation of the SEP program (2009-2010).

Table 4-6: 4th Grade, Average Hours that Teachers Spent on Class Preparation per Week

	SEP	NON SEP
2006	5.49	5.87
2009	6.87	7.75
2010	7.28	7.74
Diff 2006-2010	1.79	1.87

The results show, as advocates of accountability systems would expect, an average increase of 1.79 hours between the years 2006 and 2010 in terms of the time teachers spend preparing classes. However, Table 4-6 also shows the same trend for non-SEP schools, where in 2006 teachers spent 5.87 hours on average preparing class and where in 2010, they spent on average 7.74 hours, which represents an increase of 1.87 hours.

Working Harder: Giving more Homework

A third type of response that teachers develop under high-stakes accountability systems is that they give their students more homework or harder assignments. There is no information available prior to the implementation of the SEP program; the only data available is for the period between 2011 and 2013. Table 4-7 shows the results for 4th grade.

Table 4-7: 4th Grade, How often do you Give Math Homework this Class?

	Always		Usually		Sometimes		Never	
	SEP	NON SEP	SEP	NON SEP	SEP	NON SEP	SEP	NON SEP
2011	2.2%	3.7%	40.3%	44.4%	43.1%	40.2%	14.4%	11.7%
2013	2.2%	5.1%	45.6%	51.9%	40.3%	33.8%	11.9%	9.3%
2013-2011	0.0%	1.4%	5.3%	7.5%	-2.8%	-6.4%	-2.5%	-2.5%

One can observe that SEP schools have the same percentage (2.2%) of teachers reporting that they "always" give homework to their students in 2011 and 2013. In contrast, the percentage of teachers reported that they "always" give homework to their students increased by 1.4 percent in non-SEP schools between 2011 and 2013. Furthermore, the percentage of teachers who "always" give homework to their students is higher for non-SEP schools than SEP schools in 2013.

In 2002, the teacher questionnaire also asked how often teachers gave homework to their students. But the Likert scale is different than the scale offered in 2011 or 2013

and, therefore, it is not possible to compare schools from before and after the implementation of the SEP program. However, the teacher questionnaire in 2002 includes the option of "never" as a response when teachers are asked if they give homework to their students. This answer can be compared to similar answers given in 2011 and 2013.

Table 4-8: 4th Grade, Percentage of Teachers that "Never" Give Homework to their Students

	Never	
Years	SEP	NON SEP
2002	1.9%	4.2%
2011	14.4%	11.7%
2012	14%	10.2%
2013	11.9%	9.3%
Diff 2002-2013	10%	5.1%

Table 4-8 shows that in 2002, the percentage of teachers who indicated that they "never" gave homework to their students was 1.9 percent while teachers in non-SEP schools reported 4.2 percent. After the implementation of the SEP program in 2008, the percentage of teachers who "never" gave homework to their students in SEP schools was 11.9 percent and 9.3 percent for teachers in non-SEP schools in 2013. In brief, in 2002 the percentage of teachers from SEP schools that "never" gave homework was less than half that of non-SEP schools, but in 2013, SEP schools exceeded non-SEP schools by more than 2 percent. After the implementation of the high-stakes accountability system in Chile, the opposite would be expected: that under the pressure to improve student learning, teachers would give more homework to their students. The opposite is observed in Table 4-7 and Table 4-8.

In conclusion, using the administrative data provided by the Chilean government, it is not clear that teachers are more commonly using academic strategies following the implementation of the SEP program. In the next section, I use 23 semi-structured interviews with academics, policy makers, and teachers to explore the difficulties and

constraints that schools face in using academic strategies as a way to improve student learning.

Semi-structured interviews

I interviewed a top-level politician in charge of the implementation of the high-stakes accountability system in his downtown office at the Ministry of Education in the city of Santiago. He explained the government's expectation as to how the new accountability system will impact methods of teaching and student learning.

> We need to develop the capacity of the system to allow schools to use the information provides from the SIMCE tests to be used as powerful instruments that can be used by teachers to reflect on the pedagogical strategies, and pedagogical practices.

In fact, the assumptions and expectations under the accountability system implemented in Chile are that ranking schools and making public these ranking would help schools. This has been part of the narrative of the neoliberal prospective (Fontaine & Eyzaguirre, 2001). However, these assumptions have been criticized even by actors who have advocated for the introduction of an accountability system in the Chilean education. For example, the education director of Fundación Chile, an important think tank in Chile, she explained that they support and promote the implementation of neoliberal policies:

> The Chilean government has assumed that standardized tests are going to improve quality; they suppose that the information itself is what is going to generate the improvement. And what it improves the quality is the processes: processes of evaluation, of accountability…

As is pointed out above, one of the important critiques of Chilean accountability is that the system puts pressure on the outcome of standardized tests but does not apply enough effort or attention to the way that schools and school administrators use the results of standardized tests as a way of improving their methods of teaching, educational

projects, etc. In fact, based on the responses of teachers to the 4th grade teacher questionnaire, in 2012 only 86 percent of the 4th grade teachers said they had access to a report submitted by the Ministry of Education about the school's performance on SIMCE tests. To this, the education director of Fundación Chile said:

> In my opinion, when one starts to operationalize the information, like the SIMCE does, but doesn't allow it to trickle down [to the teachers], it doesn't go back to the instrument so it can be an effective input in a decision-making process.

The way that the accountability system is structured in Chile does not allow to schools to use the information from SIMCE tests as a means of improving their work. Instead, the central government is focused on increasing pressure on schools with the expectation that this pressure and possible sanctions will motivate schools to improve the quality of their education. In this sense, this could be one reason why it is unclear in the teacher questionnaire that SEP schools are using more academic strategies than before the implementation of the high-stakes accountability system.

Furthermore, I interviewed the director of the school of education of Universidad de Concepción. I was interested to know whether or not teachers from different geographical locations would have a different perception of the high-stakes accountability system, in particular how schools are responding to the pressure of showing better results on standardized test:

> The improvement plans are associated with making schools produce [results] through standardized tests. The problem is that the test doesn't reveal what students have learned, but rather it focuses on the results of a test. It's a measure that serves us, standardized tests are of interest to us, but they are just one aspect of the work done at a school.

Many people that were interviewed mentioned that even though the Ministry of Education is using cognitive and non-cognitive elements to create the schools rankings. Schools know that what really matters to the Ministry of Education is student performance on the SIMCE test; the rest of the elements are less relevant or secondary. In an interview, a principal from a religious private-voucher school explained that they enroll students from low-income families in Santiago and do not select their student body, which is the case with many other religious private-voucher schools in Chile. I asked him how they navigate the requirements of high-stakes accountability system and the values promoted by religious schools.

The Adventist schools arise as an answer to the concern about principles; the Christian ideals. As such, they are strong with respect to values, because they teach respect, obedience, and they teach a lot of things that are not measured in the SIMCE. The SIMCE only measures academic results, so then in this sense we are impaired. It might be true that our scores have been better, they have been rising, but it also means that we have to give up doing several other things in order to achieve this.

As the principal pointed out, under the high-stakes accountability system, schools need to be focused on elements that are directly related to standardized tests while everything else seems to be secondary because if schools do not meet the standards requested by the central government, they can receive economic sanctions or eventually be closed. However, the feedback that schools receive is not a tool that can be used by schools to review and improve their work, as was mentioned above. In this scenario, this is how a 4th grade teacher responded when asked 'What do you think would happen when the high-stake accountability system is mandatory at all schools?':

I believe that in this scenario the issue is going to be more complex, they are going to be under pressure to raise their sores but they also will be threatened with closure. So, they

are going to try to get results if the students are capable of doing so or not. It is not possible that the goals will be achieved only through a legal way.

The teacher's answer addressed one of the main concerns with the implementation of high-stakes accountability system: that schools may take a different path – different than those expected from advocates of accountability system – to improve standardized test scores without necessarily improving student learning. These behaviors have been described by Koretz (2008) as coaching, also known as "teaching to the test" or cheating. Another known behavior implies the exclusion of students or the classification of students with special needs, disciplinary sanctions, etc. These are types of behaviors that in theory the Chilean Ministry of Education should try to avoid, but the high-stakes accountability system is set up in such a way that these responses, which should be promoted as academic strategies to improve students learning, are disregarded under the strong pressure for results and the weak support and feedback that schools receive from the central government.

In Chapter 5, I will analyzed whether or not the exclusion of low-achievement students increased after the implementation of the high-stakes accountability system as a way for schools to inflate SIMCE test scores. However, as the work of Koretz has shown, there are additional strategies that schools could use, such as reallocation and alignment that would improve students' performance on standardized tests. Reallocation refers to shifting resources, including time, to emphasize the subjects and type of questions that will appear on the test, while alignment consists of matching the curriculum more closely to the material covered on the test (Koretz, 2003). In an interview with an elementary teacher about whether or not schools are using reallocation and alignment under the accountability system, he said:

Schools invest all their money in tutoring classes, it is a outsourced service; teachers from the outside work with the students after their classes, sometimes on Saturday, even during the winter vacations students received some tutoring classes. All these topics that are not math or reading are irrelevant. For example, the social sciences, art or music -- they don't give tutoring for these subjects.

This answer is consistent with many others that I have heard during my interviews; reallocation seems to be a common practice among schools, while the intensity levels vary. Even so using data provided by the Chilean government, it is not possible to quantify whether or not its use or intensity increased after the implementation of the accountability system. However, using the teacher questionnaire, I am able to identify some patterns of alignment after the implementation of SEP program (2008). Using data from 2006 and 2010, I compared three items in math curriculum that are not directly related to those items ask in the SIMCE tests. The teacher questionnaire asked to math teachers in 4th grade how often they review these items in their math class, and the results are displayed in Table 4-9.

Table 4-9: Percentage of Teachers that Reviewed Items in their Math Classes that are not Measured in the Math SIMCE Test, Before and After the Implementation of SEP Program.

		Yes		No	
		SEP	NON SEP	SEP	NON SEP
Item 1: Communication of procedures used to solve problems and the results obtained	2006	29.8%	33.0%	0.2%	0.0%
	2010	27.4%	41.8%	0.4%	0.7%
Item 2: Decision-making processes in relation to how a problem is solved, developed or modified in case the solution is not correct	2006	18.6%	28.7%	1.8%	1.3%
	2010	16.8%	28.7%	1.2%	1.0%
Item 3: Review whether or not a result obtained is correct in relation to the context of the problem	2006	24.3%	27.6%	0.5%	0.2%
	2010	20.4%	30.8%	1.3%	0.8%

Table 4-9 shows a reduction in the percentage of teachers in SEP schools that reported to cover items not related with SIMCE test. Thus, in 2006 29.8 percent of

teachers said they "completely" taught Item 1, and the percentage decreased to a 27.4 percent in 2010. In contrast, in non-SEP schools the percent of teachers who reported that they "completely" taught Item1 increased from 33 percent in 2006 to 42 percent in 2010. In relation to Item 2, SEP schools reduced the percentage of teachers who reviewed Item 2 with their students from 18.6 percent in 2006 to 16.8 in 2010 and non-SEP schools did not show any fluctuation (28.7 percent). Finally, teachers from SEP schools reduced the percentage of teachers that "completely" reviewed Item 3 from 24.3 percent (2006) to 20.3 percent (2010). In contrast, non-SEP schools increased from 27.6 percent in 2006 to 30.8 percent in 2010. In summary, Table 4-9 shows the opposite trends for SEP and non-SEP schools: while SEP schools reduced the percentage of teachers who reported to review in their classes those items on the math curriculum that are not measure directly by math SIMCE test, non-SEP schools did the opposite. I argue that this data presents some evidence of alignment among SEP schools due to the necessity of focusing on those items that will be tested by SIMCE tests.

Another aspect that remains consistent throughout all the interviews is the notion that teachers are more stressed after the implementation of the high-stakes accountability system. This mainly occurs, as teachers explained, because they need to deal with increasing levels of pressure and control from the Ministry of Education. For example, one teacher said: "Taking the SIMCE test is stressful for the majority of teachers, because school is measured based on test scores". Teachers are more conscious of the relevance of the standardized tests under the new regime, and there is an upward trend from 2000 to 2010 in terms of the relevance of the SIMCE test for teachers. For example, the percentage of teachers who give high importance to the test increased from 11 percent in 2000 to 35 percent in 2010 (CIDE, 2010). I interviewed a high-level policy maker and academic who acted as an auditor of previous Chilean governments in the area

of education, and when I asked about the problems he identified under the new accountability system, he said:

> The major problem is that you want to transform the system and at the same time help teachers using market-based instruments. The system collapses because the rules of the market oblige a school to use a group of strategies that aren't necessarily the best from a pedagogical standpoint. Because there's money involved, the schools aren't competing purely on their honor—they're competing for their survival, and the teachers are competing for their salary.

As is presented above, schools are evaluated through standardized tests that provide them with information about students learning, and they can use this information to improve their methods of teaching, school management, etc. But at the same time, the standardized tests also provide information that serve the control and regulation of the central government, which in worst case scenarios could close schools based on poor SIMCE tests performance.

One of the questions that I have tried to address in this chapter is how schools deal with this dual aspect of the accountability system, when standardized tests can be a source of information used to improve a school's work, but also the main input used to punish schools. In this sense, when I asked a professor from Universidad de Concepción of school of education, how schools respond to the pressure to achieve certain standard and thresholds on the SIMCE tests, he said:

> At the end of the day people fake results, schools are more worried about what they're going to look like in the picture and the picture is the SIMCE test. Schools show the government that they are doing things. The principals are astute and know the system and at the end of the day they are more worried about following the model [do what the government expected from schools] than from the schools by itself.

As this professor point out, because schools are under pressure for results, they learn how to game the system, and teachers and principals learn to give the Ministry of

Education what they are asking for. Falabella and Opazo (2014) conducted a qualitative study about the impact of the high-stakes accountability system on the school administration and found that the accountability system gives higher levels of structure and formalization to the school management and that teaching work under such a system is "driven toward goals." But at the same time, the accountability system creates incentives for an instrumental rationality in schools that prioritizes the achievement of goals and satisfies the formal requirements put forth by the Ministry of Education while disregarding or relegating to a second priority an autonomous professional analysis about the schools' needs and a decision-making process, contextualized in the reality of schools, that could serve to improve their work (Falabella & Opazo, 2014).

I explain that because teachers and schools administrators are not necessarily prepared to deal with a system that pushes them, because teachers do not necessarily have the skills to do what the accountability system is asking of them. In this sense, an academic told me:

> The high stakes accountability system is going to create tension among teachers, and there is a weakness here because they do not do what they need to do. The challenge is a pedagogical one, and it is how to teach. In consequence, there are capabilities that must be developed, integrated group of work, decision-making process. It is not simple; creating learning communities, what teachers mostly ask for is space to work with their colleagues.

I use the data reported in the 4th grade teacher questionnaire in (2010) to explore the differences between teachers from SEP and non-SEP schools in terms of their training.

Table 4-10: What Type of Institution 4th Grade Teachers Attended

	University		Part-Time		e-learning	
	SEP	NON SEP	SEP	NON SEP	SEP	NON SEP
2010	86.7%	95.3%	11.1%	3.4%	2.2%	1.4%

Table 4-10 shows that around 13 percent of teachers from SEP schools have been educated under special programs, in which students need to attend intensive classes

during the weekend, usually on Saturday, or by way of e-leaning or distance programs in which students use internet to access their classes and materials. These programs do not require an admissions test and generally offer low levels of quality (Ruffinelli & Rojas, 2007).

Another consideration is that the majority of the teachers who are being educated under part-time or e-learning programs are working in public schools, which concentrate students from low-income and the most disadvantaged families. In fact, according to the teacher questionnaire from 2010, 4th grade teachers who studied in part-time or e-learning programs represent 15 percent of the teaching body in public schools, while they represent only 8 percent in private-voucher schools.

One new requirement under the accountability system that has been a challenge for schools is the development of cooperative work among teachers and the development of learning communities within schools. A math teacher commented that under the increasing pressure to achieve better test scores, stress among teachers and principals makes it difficult for them to work cooperatively with their colleges

> Many of them are pressured [teachers], the SIMCE creates anxiety and stress. It is an unhealthy thing within schools, because it creates a division among teachers; you have teachers on one side and principals on the other, attacking each other. There isn't spirit of unity to work together toward the same objective.

The teacher questionnaire in 2002 and 2012 asked 4th grade teachers if they would like to move to a different school, 12.3 percent of teachers (1,117 over 9,061) said "yes" in 2002 and ten years later, after the implementation of high-stakes accountability system, 18.2 percent (1,855 of 10,064) said they would like to work in a different school. It is hard to know why more teachers are unsatisfied in their current job, but I argue – based on my conversation with teachers and academics – that although the Chilean

educational system has become increasingly demanding, it does not give teachers and schools administrations the additional resources or support to achieve the goals imposed by the central government.

It is important to point out that the SEP program not only significantly increases resources but also school autonomy in spending. However, there exists strong criticism among academics about the way schools and principals are spending the resources provided by SEP program, as much of the spending is not necessarily linked with the goal of improving the quality of education of their student body. I interviewed an assistant dean of the school of education of Pontificia Universidad Católica (PUC), one of the most important universities in Chile; he has also worked for the Ministry of Education in related positions. I asked for his opinion about how principals are using resources given by the SEP program preferential voucher. He said:

> There's a lot of money for schools [SEP schools], but they don't budget well how they use it - it's like a cash register that the school uses for very important things or to hire teachers. So then the objective was that the spending [of the resources from SEP program] was associated to improvement on quality, but this objective is not being achieved. It is very disperse [spending] and that affects the impact of the SEP program because it would be different if the money were to be spent in a more systematic way, with planning, monitoring in relation to the education plan of improvement of schools. But this is not happening and more resources do not help.

From the interviews, I have gathered the impression that that schools are not necessarily investing their economic resources in improving aspects of education related with students learning. Instead, I have the perception that principals are buying "things," but not necessarily helping schools with their educational project. For example, another academic speaking about the results of a study she conducted about SEP schools, said the following:

I believe the money is relatively poorly invested, because the first money from the SEP program was a lot money that arrived to schools and it is easy to spend when you don't have a clear program of improvement. The first thing schools did was to buy TV, interactive boards, anything, because the education project was not clear.

As mentioned above, the general perception that I have after finishing my interviews is that schools and municipalities do not have the skills to administrate the new resources or to orient their spending towards the needs and requirements of their educational project and with the final objective of improving the students learning, especially those schools with a high number of low-income families, which are the main target of the SEP program.

CONCLUSIONS

The Preferential Voucher Program (SEP) system was implemented in 2008 as a result of the pressure applied by striking students in 2006, who demanded improved quality and equity in the educational system (Salinas & Fraser, 2012).The SEP program was mainly designed to confront educational inequalities of student outcomes in the Chilean education system and for the first time, the Chilean government explicitly recognized that it is more expensive to educate students from low-income families. Therefore, the SEP program gives a preferential voucher to those schools that enroll students from low socioeconomic status. The SEP program also established a high-stakes accountability system (Murnane, Page, & Vegas, 2010). At the school level, one of the main objectives of the SEP program is the improvement of teaching methods, principals' leadership, and the overall management of the education system (Weinstein, Fuenzalida, & Muñoz, 2009). Following the typology developed by Daniel Koretz, on the possible responses of teachers under a high-stakes accountability system, and in light of the goals

defined by the SEP program, this chapter analyzes whether or not teachers of SEP schools are finding better methods of teaching, teaching more, and working harder (academic strategies) as a way to improve student learning. Furthermore, by analyzing semi-structured interviews, this chapter provides an overview of the difficulties and constraints that the current Chilean educational system faces in encouraging teachers and schools administrators to use academic strategies.

First, one type of response by teachers under the high-stakes accountability system is that teachers would work more effectively, finding better methods of teaching, for example. My findings show that in terms of feedback and the type of evaluations that teachers are using, there are similar patterns for both SEP and non-SEP schools, but that SEP schools show a larger increase in the percentage of teachers that "always" give feedback to their students on study guides and homework. In term of teaching methods, the data shows that after the implementation of the high-stakes accountability system, teachers from SEP schools decreased their use of methods that focused on individual and group work with students. After 2008, there is not a clear or unique method of teaching that SEP schools appear to be using. This could be positive, as it could show that teachers are more flexible in using different methods to teach after the implementation of accountability system, but it could also mean that there is not a clear guideline as to what teaching methods should be used by teachers from SEP schools or how they should be focused in order to ensure better and deeper learning among their students.

A second type of expected response is that teachers from SEP schools would simply teach more, spending more time overall with their students. I tested this assumption, using data about the average time that teachers spent preparing their classes before and after the implementation of SEP program. The data shows that from 2006 to 2010, teachers from SEP schools increased their preparation time by 1.79 hours, which is

in the expected direction, but non-SEP schools also showed an average increase of 1.87 hours after the implementation the accountability system.

A third type of academic strategy developed by teachers after the implementation of SEP program is that teachers would work harder, giving more homework or harder assignments to their students, for example. I used the teacher questionnaire data to analyze if the percentage of teachers in SEP that "always" give homework to their students had increased, or if the percentage of them who "never" give homework had been reduced. The results show that from 2011 to 2013, the percentage of teachers that reported to "always" give homework to their students is the same (2.2 %), but that teachers from non-SEP schools increased from 3.7 to 5.1 percent in the same period. Furthermore, using data from before the implementation of the high-stakes accountability system, the percent of teachers from SEP schools that reported to "never" give homework to their students increased from 1.9 percent in 2002 to 11.9 percent in 2013, while non-SEP schools showed an upward trend from 4.2 percent in 2002 to 9.3 percent in 2013. In brief, between the years 2002 and 2013, the percentage of teachers who "never" give homework to their students increased for SEP schools by 10 percent and only by 5.1 percent for non-SEP schools.

Finally, using 23 semi-structured interviews, I explored the difficulties and constraints that schools face in increasing the use of academic strategies as a way to improve student learning. The general conclusion, based on the content analysis of the interviews, is that the high-stakes accountability system plays a dual function. On one hand, the students' performance on SIMCE tests should be the main source used to identify which aspects of schools need to be improved, and to help the Ministry of Education to recognize areas in which they need to better support schools. One the other hand, standardized tests are at the same time use to rank schools, and if schools do not

meet the standards defined by the central government, schools can receive economic sanctions or be closed. In this fight between this two antagonist objectives, the pressure to produce results and avoid sanctions is stronger than the idea that SIMCE test could be used as a source of help or support for schools.

In this scenario, schools are focused on improving their performance on the SIMCE tests and avoiding sanctions from the central government, using a combination of non-academic strategies such as exclusion of low-achieving students from the test-taking pool as well as shifting resources, including time, to emphasize the subjects and types of questions on the test (reallocation), matching the curriculum more closely to the material covered on the test (alignment) and teaching to the test (coaching). One of the main limitations that SEP schools face in improving their use of academic strategies is the quality of the teacher training. For example, around 13 percent of teachers from SEP schools have been educated under special programs that have low levels of quality. In addition, the frequent rotation of teachers does not allow to schools to develop educational projects in the long run while the lack a leadership among principals means that resources are not used to promote educational project at their schools. These factors, combined with the need to produce results, push teachers and school administrators to seek strategies that produce results in the short run and that can be easily implemented; academic strategies do not offer these characteristics.

References

Benware, C., & Deci, E. (1984). Quality of Learning With an Active Versus Passive Motivational Set. *American Educational Research Association*, *21*(4), 755–765.

Burgess, S., & Ratto, M. (2003). *The Role of Incentives in the Public Sector: Issues and Evidence* (No. 03). Bristol.

Cameron, J., & Pierce, W. D. (1994). Reinforcement , Reward , and Intrinsic Motivation: A Meta-Analysis. *American Educational Research Association*, *64*(3), 363–423.

Chapman, D. W., Wesley, C., & Jr, S. (2000). Can High Stakes National Testing Improve Instruction: Reexamining Conventional Wisdom. *International Journal of Education Development*, *20*, 457–474.

Correa, J., Parro, D., Reyes, L., & Ugarte, G. (2012). *El impacto del voucher diferenciado en los resultados académicos: el caso chileno*.

Cullen, B. J., & Reback, R. (2006). *Tinkering Toward Accolades: Schools Gaming Under a Performance Accountability System* (No. 12286). Retrieved from www.nber.or/papers/w12286

Deci, E., Nezlek, J., & Sheinman, L. (1981). Characteristics of the rewarder and intrinsic motivation of the rewardee. *Journal of Personality and Social Psychology*, *40*(1), 1–10. doi:10.1037//0022-3514.40.1.1

Deci, E., Ryan, R., & Koestner, R. (1999). A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation. *Psychological Bulletin*, *125*(6).

Deci, E., Ryan, R., & Williams, G. (1996). Need satisfaction and the self-regulation of learning. *Learning and Individual Differences*, *8*(3), 165–183. doi:10.1016/S1041-6080(96)90013-8

Deci, E., Schwartz, A., Sheinman, L., & Ryan, R. (1981). An instrument to assess adults' orientations toward control versus autonomy with children: Reflections on intrinsic motivation and perceived competence. *Journal of Educational Psychology*, *73*(5), 642–650. doi:10.1037//0022-0663.73.5.642

Dixit, A. (1997). Power of Incentives in Private versus Public Organizations. *The American Economic Review*, *87*(2), 378–382.

Dixit, A. (2002). Incentives and Organizations in the Public Sector An Interpretative. *Journal of Human Resources*, *37*(4), 696–727.

Eyzaguirre, B., & Fontaine, L. (1999). *¿Qué Mide Realmente el Simce?* (No. 75) (Vol. 75). Santiago, Chile. Retrieved from http://www.praxis.cl/sanisidoro/Tareas/02 Actividad de lectura/Qué mide realmente el SIMCE Eyzaguirre.pdf

Falabella, A. (2014). The Performing School: The Effects of Market and Accountability Policies. *Education Policy Analysis Archives*, *22*(70), 1–29.

Falabella, A., & Opazo, C. (2014). *Sistema de Aseguramiento de la Calidad y procesos de mejoramiento: una mirada desde la gestión educativa*. Santiago, Chile. Retrieved from http://www.academia.edu/8568947/Sistema_de_Aseguramiento_de_la_Calidad_y_procesos_de_mejoramiento_una_mirada_desde_la_gesti?n_educativa

Fehr, E., & Falk, A. (2002). Psychological Foundations of Incentives. *European Economic Review*, *46*, 687–724.

Figlio, D., & Getzler, L. (2002). *Accountability, Ability and Disability: Gaming The System* (No. 9307). Cambridge, MA.

Fontaine, L., & Eyzaguirre, B. (2001). Una Estructura que Presione a las Escuelas a Hacerlo Bien *. In *¿Qué Hacer Ahora? Propuestas para el Desarrollo.* (pp. 66–111). Santiago, Chile: Centro de Estudios Publicos.

Haertel, E., & Herman, J. (2005). *A Historical Perspective on Validity Arguments for Accountability Testing* (Vol. 1522). Loas Angeles, CA.

Hamilton, L. S., Stecher, B. M., & Klein, S. P. (2002). *Test-Based Accountability in Education*. Santa Monica, CA: RAND Ecation.

Hanushek, E., & Raymond, M. (2005). Does school accountability lead to improved student performance? *Journal of Policy Analysis and Management*, *24*(2), 297–327.

Hout, M., & Stuart, E. (2011). *Incetives and Test-Based Accountability in Education*. Retrieved from http://standardizedtests.procon.org/sourcefiles/incentives-and-test-based-accountability-in-education.pdf

Jacob, B. (2007). *Test-Based Accountability and Students Achievement: An Investigation of Differential Performance on NAEP and State Assessments* (No. 12817). Cambridge, MA.

Kim, S. (2004). Accountability and School Reform in the U . S . Public School System. *KEDI Journal of Educational Policy*, *1*(1), 2–35.

Koretz, D. (2003). Using Multiple Measures to Address Perverse Incentives and Score Inflation. *Educational Measurement: Issues and Practice*, *22*(2), 18–26.

Koretz, D. (2008a). Further Steps Toward the Development of an Accountability-Oriented Science of Measurement. In K. Ryan & L. Shepard (Eds.), *The Future of Test-Based Educational Accountability* (pp. 71–91). Taylor and Francias, Hoboken.

Koretz, D. (2008b). *Measuring Up: What Educational Testing Really Tells Us*. Harvard University Press.

Koretz, D., & McCaffrey, D. (2001). *Toward a Framework for Validating Gains Under High-Stakes Conditions* (Vol. 1522). Los Angeles, CA.

Laffont, J., & Martimort, D. (2001). *The Theory of Incentives: The Principal-Agent Model*. Princeton University Press.

Manzi, J., Bogolasky, F., Grau, V., Guitiérrez, G., & Volante, P. (2014). *Análisis Sobre Valoraciones, Comprensión y Uso del SIMCE por parte de Directores Escolares de Establecimientos Subvencionados*. Santiago, Chile.

Mineduc. (2003). *Evaluación de Aprendizajes para una Educación de Calidad*. Santiago. Retrieved from http://www.agenciaeducacion.cl/wp-content/uploads/2013/02/Comision_Simce.pdf

Mineduc. (2008). *Manual de Uso de la Base de Datos SIMCE 2008 para 4 Basico*. Santiago, Chile.

Mineduc. (2012). *Impacto de la Ley SEP en SIMCE: Una Mirada a 4 años de su Implementacion*. Santiago, Chile. Retrieved from http://www.mineduc.cl/usuarios/bmineduc/doc/201208301544260.EVIDENCIASCEM8.pdf

Mineduc. (2013). Matricula por Estudiante. Retrieved January 2, 2015, from http://centroestudios.mineduc.cl/index.php?t=96&i=2&cc=2036&tm=2

Mizala, A. (2007). La Subvencion Preferencial. In C. Peña & J. J. Brunner (Eds.), *La reforma al sistema escolar : aportes para el debate* (p. 129). Santiago, Chile: Universidad Diego Portales.

Mizala, A., & Romaguera, P. (2001). *Regulación, incentivos y remuneraciones de los profesores en Chile*. CEA. Retrieved from http://books.google.com/books/about/Regulacion_incentivos_y_remuneraciones.html?id=Jb6RYgEACAAJ&pgis=1

Mizala, A., & Torche, F. (2013). ¿Logra la subvención escolar preferencial igualar los resultados educativos? *Espacio Publico*, *09*, 1–36. Retrieved from http://www.espaciopublico.cl/publicaciones/detalle.tpl?id=19

Mizala, A., & Urquiola, M. (2007). *School Markets: The Impact of Information Approximating Schools' Effectiveness* (No. 13676). Cambridge, MA. Retrieved from http://www.nber.org/papers/w13676

Murnane, R. J., Page, L., & Vegas, E. (2010). La Subvención Escolar Preferencial (SEP) en Chile: Examinando las tendencias previas y su primer año de implementación . Retrieved from www.dipres.gob.cl/594/articles-121390_doc_1pdf.pdf

Neal, D., & Schanzenbach, D. (2010). Left behind by design: proficiency counts and test-based accountability. *The Review of Economics and Statistics*, *92*(May), 263–283.

Nicol, D., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning : A model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199–218. doi:10.1080/03075070600572090

Perticara, M., Román, M., & Selman, J. (2013). *Rendimiento escolar y uso de los recursos de la ley SEP* (No. I-297). Retrieved from http://fen.uahurtado.cl/wp-content/uploads/2010/07/I-2971.pdf

Pyrczak, F. (1972). Objective evaluation of the quality of multiple-choice test items designed to measure comprehension of reading passages. *Reading Research Quarterly*, *8*(1), 62–71. doi:10.2307/746981

Romaguera, P., & Gallegos, S. (2010). Financiando la Educación de Grupos Vulnerables: La Subvencion Escolar Preferencial. In O. Larrañaga & D. Contreras (Eds.), *Las Nuevas Politicas de Proteccion Social en Chile* (pp. 1–41). Santiago, Chile: PNUD.

Roman, M. (1999). *Usos Alternativos del Simce: Padres, Directivos y Docentes* (No. 5). Santiago, Chile.

Ruffinelli, A., & Rojas, M. T. (2007). *Sistematización de la oferta de Programas Especiales de pedagogía en Educación Básica de las instituciones de educación superior chilenas*. Santiago, Chile. Retrieved from

http://cide.cl/asesoria/asesoria_desarrollo_detalle_actualizacion_sistema_ofertas_programas_pedagogia.php

Salinas, D., & Fraser, P. (2012). Educational Opportunity and Contentious Politics : The 2011 Chilean Student Movement. *Berkely Review of Education*, *3*(1), 17–47.

Stecher, B. M., & Barron, S. (2001). Unintended Consequences of Test-Based Accountability When Testing in " Milepost " Grades. *Educational Assessment*, *7*(4), 259–281.

Villarroel, G. (2012). *Mejoramiento en Resultados Académicos de la Educación Básica en Chile: ¿Primeros efectos de la ley de Subvención Escolar Preferencial (SEP)?* Universidad de Chile.

Volante, L. (2007). *Evaluating Test-Based Accountability System: An International Perspective*. Stockholm, Sweden.

Waissbluth, M. (2013, July 21). Las Pervesiones del SIMCE. *La Tercera*, pp. 1–9. Santiago, Chile.

Weinstein, J., Fuenzalida, A., & Muñoz, G. (2009). *La subvencion preferencial: desde una dificil instalación hacia su institucionalizacion*. Santiago, Chile. Retrieved from http://ww2.educarchile.cl/UserFiles/P0038/File/publicaciones/La_subvencion_preferencial_desde_una_dificil_instalacion_hacia_su_institucionalizacion.pdf

Chapter 5

The Invisible Children: What Standardized Tests Left Behind in Chile

Advocates of test-based accountability systems argue that the traditional educational system does not provide enough checks and incentives to ensure that principals and teachers maximize student learning (Cullen & Reback, 2006). Using standardized tests to hold schools accountable appears to be a simple idea, students attend schools in order to learn and tests show whether or not schools are meeting learning standards (Haertel & Herman, 2005). Schools could be classified according to the level of students' achievement so that parents can make informed decision about quality of schools, and the government can reward or punish teachers or school administrators based on an objective measurement of students' learning.

Advocates believe that the implementation of standardized tests will provide students, teachers and administrators an incentive to work harder as well as help identify struggling students and schools (Jacob, 2005). However, accountability systems often set objectives for improvement that are "simply impossible to meet by legitimate means" (Madaus & Clarke, 2001). Under high-stakes, teachers or school administrators could be tempted to make inappropriate actions when there is overstated emphasis on a single indicator and their job, salary or future is defined by students' performance in a standardized test (Nichols & Berliner, 2007).

In this model, test scores are the single most important indicator of student and school success. This can be problematic, as Donald Campbell (1979) suggests "the more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social process it is indented to monitor" (D. Campbell, 1979). This idea, Campbell's Law, has

been applied in education, particularly in the context where there is high stakes attached to an indicator, such as test score, to evaluate students, teachers and schools performance. According to this law, test scores are likely to show some degree of distortion or alteration (Madaus & Clarke, 2001).

I use Campbell's law to study the impact of using a standardized test, as a single indicator, to measure the quality of education under high-stakes accountability system in Chile. It asks the following questions: (1) are schools that participate in the high-stakes accountability program more likely to exclude low-performing students from the test-taking pool in order to improve their overall scores on standardized tests? (2) How do the schools' rankings based on standardized tests change if the students excluded from the test-taking pool were able to take the standardized test?

I use Campbell's Law to explore whether or not standardized tests show some degree of alteration under high-stakes accountability system in Chile and how extensive and meaningful could be the exclusion of low-performing students from the test-taking pool and what could its impact on schools' ranking be. This study contributes to literature in this area because the majority of studies on Chile that explore the effects of test-based accountability systems have been focused on analyzing whether or not parents use information provided by standardized tests in their selection of schools (Carrasco & San Martin, 2011; F. Gallego & Hernando, 2009; A. Torche & Sapelli, 2002). However, there appears to be no academic work examining how schools use non-academic strategies to respond to the pressure of continuing improvement established by the standardized test-based system. Particularly whether or not schools exclude low-performing students from test-taking pool and how this has an impact on the schools' rankings. The findings of this study are central for the current debate in Chile about the implementation of high-stake accountability system as a way to improve the quality of the Chilean education.

Furthermore, the findings of this study are relevant because if use of standardized test measures suffer from problems of reliability, validity or fairness, educational policy decisions based upon these scores could be misguided (Cullen & Reback, 2006; Neal & Schanzenbach, 2010).

In the Chilean case, I have found that after the implementation of the high-stakes accountability system, students with the lowest GPA are more likely to not take reading tests in the 4th grade and math tests in 4th and 8th grade. Moreover, I found that in 2nd, 4th and 8th grades schools reduce their average scores in reading and in math after imputed values are used to recalculate schools' mean scores. In fact, on average schools inflate their performance around 1.3 points on the 2nd and 4th grade tests in 2012 and this meant that 104 schools that were previously classified as "insufficient" were able to move up to "intermediate" in 4th grade in 2012.

ACCOUNTABILITY SYSTEM IN EDUCATION

As the introduction of neoliberal policies in education has gained popularity, test-based accountability systems have been seen as a necessary condition for improving the quality of education. In the last decades policy efforts in the United States have been focused on implementing these types of programs (Chapman et al., 2000; Cullen & Reback, 2006; Figlio & Getzler, 2002; Hamilton et al., 2002; E. Hanushek & Raymond, 2005; Jacob, 2007; Kim, 2004; Neal & Schanzenbach, 2010; Stecher & Barron, 2001).

The assumption of advocates of the implementation of standardized-test based accountability system is that the mechanism of rewards and sanctions created by test-based accountability systems will result in encouraging teachers and school administrators to improve at their jobs and maximize students learning. In line with this perspective, raising expectations and standards should promote better opportunities for

students and narrow the academic achievement gap between low and high performing students (Haertel & Herman, 2005).

However, the empirical evidence of the impact of test-based accountability system is mixed, some studies have found that students improve their academic achievement (E. Hanushek & Raymond, 2005; Jacob, 2005; Neal & Schanzenbach, 2010). Other research has found no impact or even negative consequences of its implementation (Brian & Levitt, 2003; Kane & Staiger, 2002; Stecher & Barron, 2001; Vasquez-Heilig & Darling-Hammond, 2008; Volante, 2007). Madaus & Clarke (2001) argue that test-based accountability systems could have diverse and even unintended negative consequences, but it is hard to evaluate how extensive these negative effects are and how they could impact schools and students (Stecher & Barron, 2001). For example, teachers and school administrators could have incentives to use non-academics strategies to ensure that the test scores give them favorable results (Brian & Levitt, 2003; Nichols & Berliner, 2007).

This study will test Campbell's law in the context of high-stakes accountability system in Chile and explores whether or not it is possible to observe some degrees of alterations or distortions under the new accountability system in Chile. Numerous studies have shown that high-stakes testing could artificially inflate students' academic achievement, particularly when the stakes are too high, teachers and school could have incentives to generate fraud or cheat (Heckman et al., 2002; Kane & Staiger, 2002). In fact, the findings indicate that school administrators and teachers, are inclined to create the appearance of the improvement of students' scores on standardized tests by using a collection of non-academic practices and thus they game the system (Dworkin, 2005; Wall, 2000).

Non-academic strategies involve excluding low achieving students from the test-pool, reclassifying them as special education students, and suspending or encouraging them to stay home during the testing window (Booher-Jennings, 2005; E. Hanushek & Raymond, 2005; Kane & Staiger, 2002). Under the test-based accountability system schools and teachers face high pressure to raise test scores and therefore it is logical for schools to leave some students, especially those who do not perform well, out of the test-taking pool. And rather than to improve the overall achievement of all students, the main objective of the implementation of test-based accountability systems, schools and teachers could simply find diverse strategies to exclude the "weak" students from the test-taking pool (Hursh, 2007). If a student does not take reading, math or sciences test, his or her scores are reported as missing value, and this student's scores do not count in calculating the school's average. In some cases, if a student does not take either of the standardized tests his/her scores will not be reported as missing value; instead this student could be excluded entirely from the dataset.

Figlio (2003) shows that during the testing window, low achieving students are most likely to receive longer suspensions than high performing students from similar violations, and students who received long suspension are much more likely to miss the standardized test (Figlio, 2003). Moreover, studies have shown that schools classify low achieving students as special education in order to exclude them from the test-taking pool (Figlio & Getzler, 2002; Haney, 2000; Jacob, 2005). Under an accountability-based system there are few incentives to advocate placing attention on those students who are already proficient or have a lower chance to pass the standardized test (Neal & Schanzenbach, 2010), and therefore teachers could pay more attention and spend more time on those students who are most likely to pass the standardized test (Booher-Jennings, 2005; Dworkin, 2005; Hursh, 2007). Furthermore, schools and teachers could

react to the pressure to raise test score by decreasing time on subjects that are not part of or are less emphasized on the exams. Thus, students could spend more time working on math or reading problems while reducing the time allotted to the sciences, history, arts, foreign languages or writing because they are not topics that will be tested. This can lead to narrowing and oversimplifying the curriculum (E. Baker et al., 2010; Carnoy et al., 2000; Hamilton et al., 2002; Stecher & Barron, 2001; Winters et al., 2010).

From low to high-stakes accountability systems

The use of standardized tests has become an important part of the debate surrounding Chilean education, especially regarding whether teachers and schools administrators should be accountable for students' scores on the SIMCE test. As Mizala and Urquiola (2007) point out above, the Chilean accountability system does not hold teachers or schools administrators accountable for students' performance in standardized tests. However, the 2006 secondary student movement and 2011 university student strike created the conditions to change this. The student movement demanded the regulation of the market-driven Chilean educational system and the strengthening of public education. The uprising resulted in the resignation of two education ministers and a shift in the agenda of the Chilean Congress, which has now had to make educational reform its top priority (Goldman, 2012). As a respond to students and civil society to improve the quality of education, in 2011 the Chilean congress passed Law #20,529 that created two new institutions: Superintendency and Agency of education, the role of the superintendence of education is to supervise the administrators of the schools recognized by the Ministry of Education Act according to the norms and rules established by law. The Agency's mission is to evaluate and orientate the educational system in order to improve its quality and equity. Moreover, the Agency must rank all schools in the system

according their students' learning in standardized tests and other non cognitive indicators of quality of education (Mineduc, 2011) Within this new accountability system those schools that cannot achieve the minimum requirements defined by the Agency of Quality of Education will face consequences, the most severe one it is the closure of schools. Thus, Chile has started a new type of educational reform centered in the use of mechanism of pressure and accountability to improve quality and equity of the system, which call for a high degree of technical sophistication and ethical consideration to implement the new requirements of pressure, control and surveillance (Carrasco & San Martin, 2012)

The implementation of the Superintendency and the Agency of Quality of Education has not been without its critics. The strong emphasis that the reforms imposes in the use of standardized tests to classify, surveillance and penalize schools has particularly not been welcome in many sectors. The argument is that the SIMCE test stopped of being a standardized test to measure students' learning and it became a central aspect of the educational system and it has established a harmful rationality (Avalos, Leyton, Soto, & Garcia-Huidobro, 2013). Chile shows an imbalance between the strong emphasis on standardized tests to make schools accountable, but little effort to use it as a learning process for teachers and principals to improve their job (Manzi, 2014). However, the director of the Agency of Quality of Education pointed out that the arguments against the implementation of the high stakes test were partial and insufficient due to there is not evidence regards to the negative impact of use of standardized test in Chile (Izquierdo, 2013). In this term this study is relevant because shows empirical evidence of unintended consequences of the use of standardized test in Chile.

METHODS AND DATA

I use two different data sets about on Chile's educational system to examine these questions. Since the data come from different surveys, grades, and years, the following section explains in more detail the information and sources contained in the data set:

(a) Sistema Nacional de Evaluación de Calidad de la Educación (Educational Quality Measurement System, SIMCE). As I explained previously in Chapter 3, SIMCE scores report the performance of schools in different subject areas (Math, Spanish, and Sciences) in comparison to previous years, other schools, and other grade levels within the same school. This research relies on data from the standardized SIMCE test applied to 2nd, 4th and 8th grades from 2002 to 2013 (see table 5-1); the unit of analysis is student-level.

Table 5-1: Grades and Years when SIMCE Test was Administered

	Years								
Grades	2002	2005	2006	2007	2008	2010	2011	2012	2013
2nd								√	√
4th	√	√	√	√	√	√	√	√	√
8th				√			√		√

b) The second data source comes from the Department of Research of Chilean Ministry of Education, this is called *student achievement* (Rendimiento del Estudiante). The unit of analysis is individual, the dataset provides information on the student's final GPA obtained, rate of attendance, and whether or not the student passed or failed his/her grade during the period 2002-2013. Furthermore, it provides a student identification number that can be use to merge with data from the National Assessment of Mathematics and Spanish Achievement (Mineduc, 2013a). To ensure that the matches were correct, I merged SIMCE test and student achievement data using the unique id of school and a unique id of student and the student's gender. Schools with fewer 6 students were

excluded because their error estimation is too high to draw a valid conclusions about schools (Mineduc, 2008). In addition, students in special education called "integrated students" were excluded from the sample because, in Chile, integrated students do not take standardized tests, but are recorded as missing values in the dataset for some years.

Furthermore, for the year 2012 and 2013, the Chilean Ministry of Education reports on all students in the system whether or not they took test standardized, but from 2002-2011 only students who at least take one of the standardized tests are reported, and therefore, I have to constrain the analysis only to those students who at least take one test in order to make valid estimations from 2002-2013.

In order to explore whether or not schools use non-academic strategies to improve their performance on standardized tests, I used a logistic regression model with fixed effects at school levels from 2002 to 2013 for 2nd, 4th and 8th grades. It is important to clarify that there is no data for 2nd grade before 2012, because testing of 2nd graders only began in that year.

The dependent variable in the logistic regression is equal to 1 if the student missed the test and 0 otherwise. The independent variables are the student's final grade in his/her class, rate of attendance, family income, father's education, mother's education, a lag variable of the school's classification, and a dummy variable if schools participate in SEP program, a dummy variable for years, and a fixed effects at the school level. In Chile students with less than 85 percent of attendance or with GPA less than or equal to B- are most likely to fail their class. Table 5-2 shows the conversion between the Chilean GPA and the US GPA.

Table 5-2: Conversion from the Chilean GPA to US GPA

Chilean GPA	Grade Description	US GPA	Notes
6.50 - 7.00	Very Good	A	
6.00 - 6.49		A-	
5.50 - 5.99	Good	B+	
5.00 - 5.49		B	
4.50 - 4.99		B-	
4.00 - 4.49	Sufficient	C	Lowest passing grade
0.00 - 3.99	Insufficient	F	

Source: http://www.classbase.com/Countries/Chile/Grading-System

I run two different models for math and reading tests in 4th and 8th grade. The dependent variable is 1 if a given student within a given school did not take the standardized test and is 0 otherwise. Gender is scored as 1 for male students and 0 otherwise. The student's scores in reading or math test are separated into three dummy variables: "insufficient", "intermediate", and "advantaged." The dummy variable for students who obtained "advanced" scores is excluded from the model, this outcome serving as the baseline. SEP school is a dummy variable equal 1 if school participates in the SEP program and 0 otherwise. The student's annual attendance rate is divided into five dummy variables, and the dummy variable for the lowest annual attendance (less than 85 percent) is excluded from the analysis. The students' GPA are separated into 5 dummy variables, and the dummy variable for the lowest GPA (equivalent to an B-) is excluded from the analysis. The students' family income is split into five dummy variables, and the dummy variable for the lowest family income (less than US $400 monthly) is excluded from the analysis. Fathers' and mothers' education is divided into three dummy variables: primary, secondary and college education. The dummy variable

for primary education is excluded from the analysis. The lagged variable of school classification is separated into three dummy variables: insufficient, intermediate and advanced. The dummy variable for schools ranked as "advanced" is excluded from the analysis.

I created an interaction term between schools that participate in the high-stakes program (SEP schools) and low-performing students. In this sense, I evaluate whether SEP schools, after participating in the high-stakes accountability program, were more likely to exclude low-achievement students from the test-taking pool. Furthermore, I created a term to measure the interaction between schools that are classified as "insufficient" and low-performing students. Thus, I evaluate whether schools that exhibit low performance on SIMCE tests have higher incentives than "advanced" schools to exclude low-achievement students from the test-taking pool in order to artificially increase their scores on standardized tests.

Finally, I will work on multiple imputation models in order to estimate the missing values present on the reading and math tests. The Chilean government ranks schools according to different levels of achievement on reading and math tests and I will use the values imputed, previously excluded, to recalculate the classification of schools and explore whether or not it is possible to observe any difference after the imputed values are included. The missing values on the reading and math tests are assumed to be Missing At Random (MAR). This means that the missing values of Y, in this case reading and math, is unrelated to the value of Y itself, after controlling for other variables in the analysis (Allison, 2002). As a consequence, this study assumes that the missing

values of test scores are not related to the values of test scores by themselves, instead the missing values are linked to the final grade obtained by the students in their courses, their attendance rate, and family's and schools' characteristics. In order to impute the missing values on the reading and math tests I used SAS's Multiple Imputation (MI) routine. When this approach is used, it correctly produces estimates that are consistent and asymptotically normal when the data are MAR (Allison, 2002).

Since students are nested within schools, I imputed the data using a panel data set. One strategy to impute the missing values at student level is to create dummies variables for each school, but since in some cases the number of students per school is small, imputing student variables with school dummies can overfit the data. As a consequence, since there are no missing values at the school level, I impute student variables with school-level variables, including means of student variables. I use separated models to impute missing data in the 2nd, 4^{th} and 8th grades. See Table 5-3 for more details.

Table 5-3: Variables Used in The Multiple Imputation Models

	Label	**Measured**
Mathematics	Math	Continuous variable
Reading	Reading	Continuous variable
Sciences	Sciences	Continuous variable
Student level		
Gender	Gender	1 = men 0 = woman
GPA	GPA	Dummy variables for 5 categories
Attendance rate	Attendance	Dummy variables for 5 categories
Family level		
Family Income	Income	Dummy variables for 5 categories: Monthly income in US dollars Less $400 / $401-1,000 / $1,001-2,000 /$ 2,001-2,300 / More than $2,300
Father Education	fedu	Dummy variables for 3 categories: Primary, secondary and college education
Mother Education	medu	Dummy variables for 3 categories: Primary, secondary and college education
School level		
Average of students' GPA	avgrades	Mean of GPA at school level
Average of students' attendance	avattendance	Mean of percentages at school level
Year	years	Dummy variables for each year

RESULTS

In this section, I present the findings of two analyses. The first section explores if SEP schools increase the likelihood of excluding low-performing students from the test-taking pool after the implementation of the high-stakes accountability program. The second part examines the significance of the impact of the exclusion of the low-achieving students from the test-taking pool on the schools' ranking.

Using a logistic regression model with fixed effectss at the school level, and after controlling for student attendance rates, as well as their family and school characteristics, I found that low-performing students in 4th grade are more likely not to take the reading test than their peers who receive a final GPA of A when their school participates in the

SEP program. For example, when schools participate in the high-stakes accountability program (SEP), students who obtained a GPA of B- or less were 25.2 percent more likely to not have taken the reading test in 4th grade when compared to schools that do not participate in the SEP program and their peers who had finished the academic year with a GPA of A. The interaction term is not statistically significant in 8th grade. Furthermore, if a school was classified as "insufficient" during the previous year, 4th grade students with GPA of B- or less were more than 8 percent more likely to have not taken the reading test in comparison to students from "advanced" schools. Nevertheless, the interaction term is not statistically significant.

As family income increased, 4th grade students became more likely to have taken the reading test, but the coefficients are not statistically significant in 8th grade. For example, a 4th grade student from a family with a monthly income of $2,300 or higher was 18 percent less likely to miss the reading test than a student from a family with the lowest level of income. Furthermore, the overall coefficients for the mother's and father's education are not statistically significant in 4th and 8th grade. Table 5-4 shows the results of the logistic regression model.

Table 5-4: Logistic Regression Model with Fixed Effects at School Level (Odds Ratios) for Missing Values in Reading Test

	4th grade	8th grade
Student's attendance rate		
85-90%	0.7580***	0.6107***
90-95%	0.6316***	0.4448***
95-100%	0.4278***	0.2774***
Student's GPA		
B	0.8206***	0.7503***
B+	0.6458***	0.6714***
A-	0.5156***	0.5844***
A+	0.4249***	0.4873***
Income of family		
US 401-1,000	0.8988***	0.955
US $ 1,001-2,000	0.8464***	0.8422**
US $ 2001-2,300	0.7589***	0.8676
More than US $ 2,300	0.8202**	0.8386
Father Education		
Secondary	0.9586	1.0094
College	0.9772	1.017
Mother Education		
Secondary	1.0559**	0.9949
College	1.0497	1.0301
Year		
2002	Baseline	
2005	0.8964**	
2006	0.8196***	
2007	1.1249**	Baseline
2008*	1.1315**	
2010	1.0111	
2011	1.2187***	0.8999
2012	1.6288***	
2013	1.4681***	1.2031**
Interaction terms		
Sch. “insufficient” x low-achievement student	1.0818	1.0002
School classified as “insufficient”	1.0837**	0.8904*
SEP school x low-achievement student	1.2522***	1.0104
SEP school	0.9878	1.0157
Number of students	1,318,206	304,606

Legend: * p<.1; ** p<.05; *** p<.01

* Implementation of SEP program

In relation to the variable year, it is possible to observe an upward trend in the odds that students would be excluded from the test-taking pool in reading in 4th and 8th grades. For example, in 4th grade, the odds of a student being excluded from the reading test-taking pool decreases by 11 percent in 2005 in relation to 2002 (baseline), but the odds increase by 46 percent in 2013 in relation to 2002.

In brief, based on the results of the logistic regression there is evidence that since schools participating in the SEP program are more likely to exclude low-achieving students from the reading test 4th grade. However, I have not found evidence of a relationship between SEP schools and the exclusion of low-performing students in reading test in the 8th grade.

The relationship between low-performing students, SEP schools, and exclusion from the test-taking pool on the standardized math test is shown in Table 5-5. Similar to the results for the reading test, I found evidence that low-performing students in 4th and 8th grade were more likely to be excluded from the math test after the implementation of the high-stakes accountability program. For example, schools that participate in the SEP program increase their likelihood of excluding low-achievement students from the math test by 25.2 percent in 4th grade and 53.3 percent in 8th grade. Both coefficients are statistically significant at a 0.05 alpha level.

Table 5-5: Logistic Regression Model with Fixed Effects at School Level (Odds Ratios) for Missing Values in Math Test

	4th grade	8th grade
Student's attendance rate		
85-90%	0.8726***	0.7970**
90-95%	0.7500***	0.6527***
95-100%	0.6417***	0.5124***
Student's GPA		
B	0.7521***	0.8525
B+	0.5279***	0.7214*
A-	0.4112***	0.5898***
A+	0.3025***	0.4279***
Income of family		
US 401-1,000	0.9421	0.8595**
US $ 1,001-2,000	0.939	0.8872
US $ 2001-2,300	0.7596**	0.5975**
More than US $ 2,300	1.0369	0.6307**
Father Education		
Secondary	0.9944	1.044
College	1.0102	1.1211
Mother Education		
Secondary	1.0052	1.0347
College	1.0383	1.1563
Year		
2002	Baseline	
2005	0.6863***	
2006	0.4559***	
2007	0.7471***	Baseline
2008*	0.7838***	
2010	0.7491***	
2011	1.0497	1.4379**
2012	1.5135***	
2013	0.8943	1.6368***
Interaction terms		
Sch. “insufficient” x low-achievement student	1.0845	0.9565
School classified as “insufficient”	1.1005*	0.8915
SEP school x low-achievement student	1.3322***	1.5337***
SEP school	0.9343	0.7110*
Number of students	933,269	131,398

Legend: * p<.1; ** p<.05; *** p<.01

* Implementation of SEP program

The analysis shows that as students' family income increases students reduce the predicted odds of being excluded from taking the math test in 4th and 8th grade, but only one of these coefficients is statistically significant at the .05 alpha level in 4th grade.

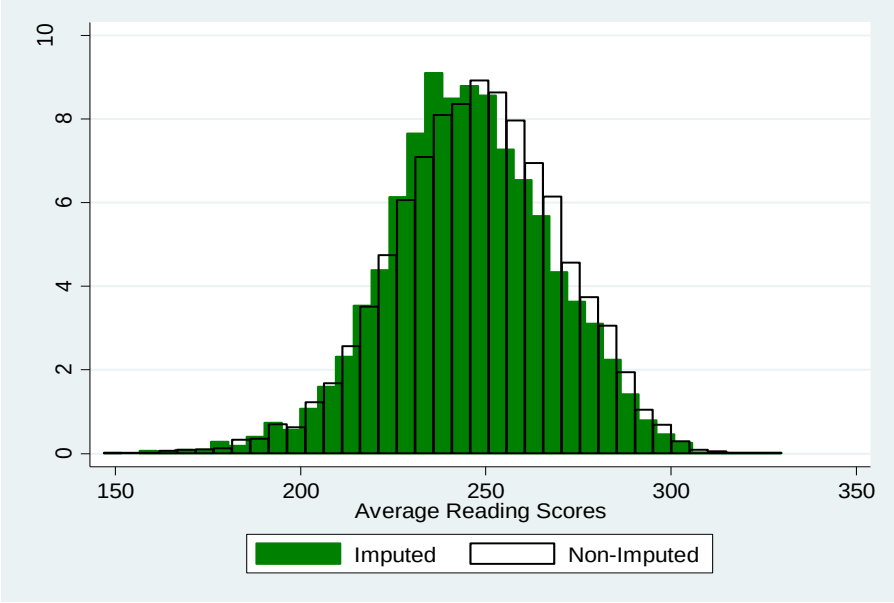
Moreover, as father's education increases, the predicted odds of students do not take the math test decrease, but none of the coefficients are statistically significant. The same case for mother's education, as mother increases their education level the predicted odds of student missing math test decrease, but none of the coefficients are statistically significant at .05 alpha level in 4th and 8th grade. In sum, there is evidence that schools that participate in the high-states accountability program (SEP program) increase the likelihood of excluding low-achieving students from the test-taking pool for the math test in 4th and 8th grade.

Finally, GPA seems to be more relevant during early levels of primary education than later ones in determining which students do or do not take the reading test. I hypothesize that this is due to the fact that as low-performing students progress, schools employ different mechanisms to exclude underperforming students from the test-taking pool. For example, schools might classify students in special education programs to exempt them from taking the test; they could drop out, or receive disciplinary sanctions during the test-taking windows.

The second research question asks whether or not the schools' rankings would change if the students excluded from the test-taking pool would be able to take the standardized test. To address this question, I use multiple imputation models in order to estimate the missing values in math and reading and recalculate the schools' rankings based on the SIMCE test. The results shows that in a reading test for 2nd grade, schools' means decline on average 1.42 points in 2012 and 1.32 points in 2013, when imputed values are included in the calculation.

The national average for reading scores for 2nd graders on the SIMCE exams, decreases from 252.7 points to 251.3 in 2012 and from 254.2 to 252.9 in 2013 after imputed values are included in the model. As we observe in figure 5-1, the distribution of scores on the reading test move to the left, which means that the overall average of schools' performance on the reading test decreases for the 2nd grade when imputed values, most of them low-performing students, are included in the school's mean. This shift to the left has implications in terms of how schools are classified by the government, which would impact the funding or technical support that they could receive from the Ministry of Education.

Figure 5-1: The Schools' Average Reading Scores Using Imputed and Non-Imputed Values in the 2nd Grade 2012

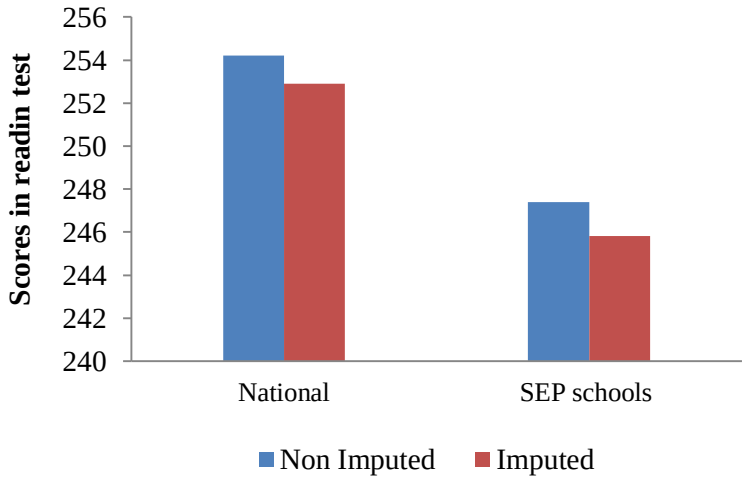


Source: Author's calculation based on Simce 2012

The differences between imputed and non-imputed data increase when schools that participate in the SEP program are compared. For example, at the national level, when imputed data is used to calculate the schools' mean on the reading test, scores

decline 1.3 points and scores go down to 1.6 points when the mean of SEP schools is calculated using imputed data. These differences are shown in Figure 5-2.

Figure 5-2: Schools' Mean on Reading Test Using Imputed and Non-Imputed Data in 2nd Grade in 2013.



Source: Author's calculation based on Simce 2013

In contrast to the 2nd grade, in 4th grade schools can be classified into three levels based on their average scores, if a school obtains an average score less than 241 points on the reading test, it is classified as "insufficient." If a school falls into the range of 242-280 points, it is defined as "intermediate," and if the school's mean on the reading test is more than 281 it is classified as "advanced." In the case of math scores, schools are defined as "insufficient" if their mean score in math is equal to or less than 232 points, "intermediate" if their mean falls in the range of 241-283 points and finally schools are defined as "advanced" if their mean scores in math is greater or equal to 284 points (Mineduc, 2007a). In this sense, a school could be classified as "insufficient" in math and "advanced" in reading; this is because reading and math tests have different thresholds to determine the schools' classification.

I used the categories elaborated for 4th grade in reading and math scores to classify schools, and as table 5-6 shows after imputed values are used to recalculate schools' averages in reading, the percentage of schools classified as "insufficient" increases 1.2 percent, those defined as "intermediate" increase 1.4 percent as well as school classified as "advanced" declines by 2.4 percent.

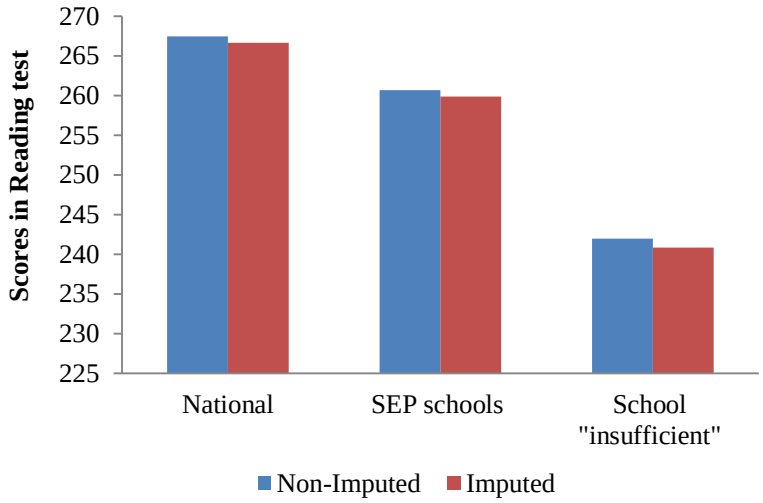
Table 5-6: Classification of Schools Based on Reading Scores at the 4nd Grade Level in 2012

Type of schools	Imputed values	Non-Imputed
Insufficient	17.0%	15.8%
Intermediate	58.2%	56.8%
Advanced	24.9%	27.3%

In fact, in 2012 the number of schools that move down from "intermediate" to "insufficient" after using the imputed data was 104 schools, and 83 schools changed their classification from "advanced" to "intermediate." This is a small number compared with the total number of the schools in the system in 4th grade, which were 5,689 schools in 2012. But it shows that if schools are willing to exclude low-performing students from reading tests can improve their classification in the Chilean system.

Figure 5-3 shows the difference between schools' mean reading scores, between non-imputed data and the multiple imputed data, increase when a school is classified the previous year as "insufficient" and also when schools are participating in the SEP program. For example, in 2013, using all schools (national level), the difference between imputed and non-imputed data in relation to the schools' mean on the reading test was 0.75 points. For those schools that participate in the high-stakes accountability program, the difference increased to 0.85 points, and finally when schools are classified as "insufficient" in the previous year, the difference between non-imputed and imputed data increased to 1.13 points on the reading test.

Figure 5-3: Schools' Mean Reading Test Using Imputed and Non-Imputed Data in 4th Grade in 2013.



Source: Author's calculation based on Simce 2013

Math tests follow the same pattern as the reading tests in that the differences between imputed and non-imputed data are used to calculate schools' mean for those schools that participate in high-stake accountability program or were classified as "insufficient" the previous year. For example, in 2013, at the national level, the difference in schools' mean on the math test was 0.82 points. The difference increased to 0.97 points for SEP schools and for those schools classified as "insufficient," the imputed data is 1.24 points lower than for non-imputed data.

The Chilean Ministry of Education classifies schools in 8th grade using similar categories to those used in the 4th grade, but different thresholds are used to determine the schools' classification. Thus, if a school obtains on average less than 235 points in reading or 276 points in math, this school is classified as "insufficient." If a school falls in the range of scoring 235-285 points on reading test or 276-320 points on the math test it is classified as "intermediate." Finally, a school is defined as "advanced" if its average is higher than 286 points in reading or 320 in math (Mineduc, 2010).

I used the categories elaborated for 8th grade in reading and math scores to classify schools, and as table 5-7 shows after imputed values are used to recalculate schools' averages in reading. The percentage of schools classified as "insufficient" increases 1.1 percent, those defined as "intermediate" decrease 0.9 percent as well as school classified as "advanced" declines by 0.3 percent.

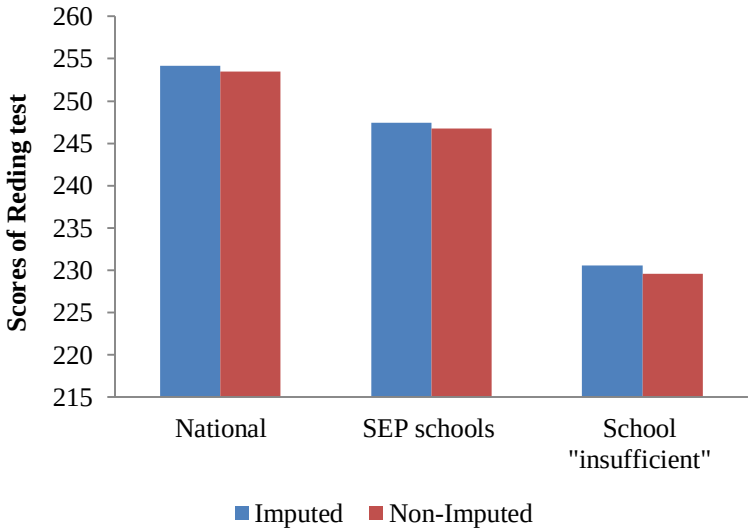
Table 5-7: Classification of Schools Based on Reading Scores at the 8nd Grade Level in 2013

Type of schools	Imputed values	Non-Imputed
Insufficient	29.5%	28.4%
Intermediate	61.4%	62.3%
Advanced	9.0%	9.3%

The number of schools that changed their status from "intermediate" to "insufficient" was 65 schools, and 15 schools that were "advanced" change to "intermediate" after imputed values were used to calculate schools' mean on the reading test.

Figure 5-4 shows the difference between schools' mean reading scores between the original data and the multiple imputed data. The data for 2013 at the national level on the reading test shows that the schools' mean the reading test decreases 0.6 points when imputed data are used. Furthermore, for those schools that participate in the high-stakes accountability program, the schools' mean the reading test declined 0.72 points when imputed data was used. And finally, among schools that are classified as "insufficient," their reading scores decline 1 point when the imputed values of students who did take the reading test are used to recalculate the schools' average.

Figure 5-4: Schools' Mean on Reading Test Using Imputed and Non-Imputed Data in 8th Grade in 2013.



Source: Author's calculation based on Simce 2013

In relation to the math test, when the imputed values of those students who did not take the math test are used to recalculate schools' classification, the results are the following. At the national level, the schools' average on the math test decreased 0.6 points when imputed data are used. This means that if those students excluded from the test-taking pool are included, national scores in math decline. Moreover, for those schools that participate in the SEP program, the schools' mean on the math test was 0.7 points lower using imputed data. When schools are classified as "insufficient," their math scores go down 1 point if imputed values are used to calculate the schools' mean.

CONCLUSIONS

This research tests Campbell's law to study the impact of using a standardized test, as a single indicator, to measure the quality of education. As Donald Campbell (1979) suggests "the more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social process it is indented to monitor" (D. Campbell, 1979).

Various studies have explored the negative impact of test-based accountability systems under high-stakes accountability systems, and even though the findings are not conclusive, their results show that high-stakes testing could have diverse and even unintended negative consequences, even though it is hard to evaluate how extensive these negative effects are and how they could impact schools and students (Stecher & Barron, 2001). I test the Campbell's law to explore whether or not standardized tests show some degree of distortion or alteration after the implementation of the high-stakes accountability system in Chile in 2008 and how extensive and meaningful could be these unintended consequences for schools and students.

I have shown that even when standardized tests and their results are mainly employed to help inform parents about school quality, and they have no direct consequences for teachers or school administrators, it possible to observe some degree of alteration or distortion. I found that low-performing students after the implementation of the SEP program were more likely to be excluded from taking the reading tests in 4th grade and in math tests in 4th and 8th grade. I argue that students' GPA play a more important role as a mechanism for identifying low-performing students at lower grades than in higher ones. I hypothesized that this is due to that as low-performing students' progress through school, there are different mechanisms used by schools to exclude low-

performing students. For example, low-performing students might be classified in special education, they could drop out or receive disciplinary sanctions during the test-taking windows, but these strategies are more difficult to use these non-academic strategies when students are just starting in the system.

Furthermore, I found evidence that in 2nd, 4th and 8th grades, schools would reduce their average scores in reading and in math after imputed values are used to recalculate schools' mean scores. For example, on the 2013 reading test, on average schools inflate their performance by 1.3 points in the 2nd grade, by 0.7 points in the 4th grade, and by 0.6 points in the 8th grade. Furthermore, using the imputed values, the percentage of schools classified by the Ministry of Education as "advanced" and "intermediate" would decrease and those defined as "insufficient" would increase.

References

Allison, P. (2002). *Missing Data* (Quatitativ.). Sage Publictaion.

Avalos, B., Leyton, M., Soto, V., & Garcia-Huidobro, J. E. (2013, September 26). Carta abierta por un Nuevo Sistema de Evaluación Educacional. *El Mostrador*, p. 6. Santiago, Chile. Retrieved from www.elmostrador.cl/media/2013/Carta-abierta.pdf

Baker, E., Barton, P., Darling-Hammond, L., Haertel, E., Ladd, H., Linn, R., … Shepard, L. (2010). *Problems with the Use of Students Test Scores to Evaluate Teachers* (No. 278) (Vol. 03). Washington, DC.

Booher-Jennings, J. (2005). Below the Bubble: "Educational Triage" and the Texas Accountability System. *American Educational Research Journal*, *42*(2), 231–268.

Brian, J., & Levitt, S. (2003). Rotten Apples: An investigation of the Prevalence and Predictors of Teacher Cheating. *The Quaterly Journal of Economics*, (August), 843–877.

Campbell, D. (1979). Assessing the impact of planned social change. *Evaluation and Program Planning*, *2*(1), 67–90.

Carnoy, M., Loeb, S., & Smith, T. L. (2000). *Do Higher State Test Scores in Texas Make For Better High School Outcomes? The Civil Right s Project*.

Carrasco, A., & San Martin, E. (2011). *Exploring New Approaches in School Choice Research: A Partial Examination of Educational Market-Driven Hypothesis From Value-Added Models*.

Carrasco, A., & San Martin, E. (2012). *Clasificacion de Escuelas en la Nueva Institucionalidad Educativa: Contribucion de Modelos de Valor Agregado para una Responsabilidad Justa* (No. 53). Santiago, Chile.

Chapman, D. W., Wesley, C., & Jr, S. (2000). Can High Stakes National Testing Improve Instruction: Reexamining Conventional Wisdom. *International Journal of Education Development*, *20*, 457–474.

Cullen, B. J., & Reback, R. (2006). *Tinkering Toward Accolades: Schools Gaming Under a Performance Accountability System* (No. 12286). Retrieved from www.nber.or/papers/w12286

Dworkin, a. G. (2005). The No Child Left Behind Act: Accountability, High-Stakes Testing, and Roles for Sociologists. *Sociology of Education*, *78*(2), 170–174.

Figlio, D. (2003). *Testing, Crimen and Punishment* (No. 11194). New York City. Retrieved from http://www.nber.org/papers/w11194

Figlio, D., & Getzler, L. (2002). *Accountability, Ability and Disability: Gaming The System* (No. 9307). Cambridge, MA.

Gallego, F., & Hernando, A. (2009). *School Choice in Chile: Looking at The Demand Side* (No. 356). Retrieved from http://www.economia.puc.cl/docs/dt_356.pdf

Goldman, F. (2012, April 22). Camila Vallejo, the World's Most Glamorous Revolutionary. *New York Times*. Retrieved from http://www.nytimes.com/2012/04/08/magazine/camila-vallejo-the-worlds-most-glamorous-revolutionary.html?pagewanted=all&_r=0

Haertel, E., & Herman, J. (2005). *A Historical Perspective on Validity Arguments for Accountability Testing* (Vol. 1522). Loas Angeles, CA.

Hamilton, L. S., Stecher, B. M., & Klein, S. P. (2002). *Test-Based Accountability in Education*. Santa Monica, CA: RAND Ecation.

Haney, W. (2000). The Myth of the Texas Miracle in Education. *Education Analysis Policy Archives*, *8*(41). Retrieved from http://epaa.asu.edu/ojs/article/view/432

Hanushek, E., & Raymond, M. (2005). Does school accountability lead to improved student performance? *Journal of Policy Analysis and Management*, *24*(2), 297–327.

Heckman, J. J., Heinrich, C., & Smith, J. (2002). *The Performance of Performance Standards* (No. 9002). Cambridge, MA.

Hursh, D. (2007). Assessing No Child Left Behind and the Rise of Neoliberal Education Policies. *American Educational Research Journal*, *44*(3), 493–518.

Izquierdo, S. (2013, October 21). ¿Eliminar el Simce? *La Tercera*, pp. 9–10. Santiago, Chile. Retrieved from www.latercera.cl

Jacob, B. (2005). Accountability, incentives and behavior: the impact of high-stakes testing in the Chicago Public Schools. *Journal of Public Economics*, *89*(5-6), 761–796.

Jacob, B. (2007). *Test-Based Accountability and Students Achievement: An Investigation of Differential Performance on NAEP and State Assessments* (No. 12817). Cambridge, MA.

Kane, T. J., & Staiger, D. O. (2002). The Promise and Pitfalls of Using Imprecise School Accountability Measures. *Journal of Economic Perspectives*, *16*(4), 91–114.

Kim, S. (2004). Accountability and School Reform in the U . S . Public School System. *KEDI Journal of Educational Policy*, *1*(1), 2–35.

Madaus, G., & Clarke, M. (2001). The adverse impact of high stakes testing on minority students: evidence from 100 years of test data. In G. Orfield & M. Kornhaber (Eds.), *Raising standars or raising barriers? Inequality and high stakes testing in public education* (p. 51). The Century Foundation. Retrieved from http://files.eric.ed.gov/fulltext/ED450183.pdf

Manzi, J. (2014, March 17). ¿Ganamos algo eliminando el Simce? *La Tercera*, pp. 2–3. Santiago, Chile. Retrieved from www.latercera.cl

Mineduc. (2007). *Aplicacion de la Metodologia Para Establecer Puntajes de Corte en las Pruebas de Corte en las Pruebas SIMCE 4 Basico*. Santiago, Chile.

Mineduc. (2008). *Manual de Uso de la Base de Datos SIMCE 2008 para 4 Basico*. Santiago, Chile.

Mineduc. (2009). *Primer Reporte de Resultados Nacionales SIMCE*. Santiago, Chile.

Mineduc. Ley 20.529: Sistema Nacional de Aseguramineto de la Calidad de la Educacion Parvularia, Basica y Media y su Fiscalizacion (2011). Valparaiso. Retrieved from http://www.cned.cl/public/secciones/seccionsnac/normativa/LEY-20529_27-AGO-2011.pdf

Mineduc. (2013). *Base de Datos de Rendimiento Escolar 2002-2012*. Santiago, Chile.

Neal, D., & Schanzenbach, D. (2010). Left behind by design: proficiency counts and test-based accountability. *The Review of Economics and Statistics*, *92*(May), 263–283.

Nichols, S. L., & Berliner, D. C. (2007). Collateral damage: how high-stakes testing corrupts America's schools. Cambridge, Mass.: Harvard Education Press. Retrieved from http://catalog.hathitrust.org/Record/005555685

Stecher, B. M., & Barron, S. (2001). Unintended Consequences of Test-Based Accountability When Testing in " Milepost " Grades. *Educational Assessment*, *7*(4), 259–281.

Torche, A., & Sapelli, C. (2002). *Subsidios al Alumno o a la Escuela: Effects en la Eleccion de Colegios Publicos* (No. 202).

Vasquez-Heilig, J., & Darling-Hammond, L. (2008). Accountability Texas-Style: The Progress and Learning of Urban Minority Students in a High-Stakes Testing Context. *Educational Evaluation and Policy Analysis*, *30*(2), 75–110.

Volante, L. (2007). *Evaluating Test-Based Accountability System: An International Perspective*. Stockholm, Sweden.

Wall, D. (2000). The impact of high-stakes testing on teaching and learning : can this be predicted or controlled? *System*, *28*, 499–509.

Winters, M. a., Trivitt, J. R., & Greene, J. P. (2010). The impact of high-stakes testing on student proficiency in low-stakes subjects: Evidence from Florida's elementary science exam. *Economics of Education Review*, *29*(1), 138–146.

Chapter 6: Conclusions

This dissertation has evaluated the macro-level impacts of market-oriented reforms in education in Chile. In particular, it has examined the intended and unintended consequences of the introduction of competition and the high-stakes accountability system to the Chilean educational system.

Influenced by Milton Friedman and under the military regime of Augusto Pinochet, Chile implemented a universal voucher educational voucher program in 1981. The new system was set up under the premise that the role of the state in education should be reduced and that the government monopoly would be replaced by free market competitors. The democratic government took office in 1990, and even though it was in disagreement with a number of military regime's educational policies, it did not introduce any substantive changes to the model. In fact, in 1981 private-voucher schools enrolled 15.1 percent of total students in primary and secondary education and in 2013, more than half of students (55 percent) attended private-voucher schools. Furthermore, as part of the neoliberal agenda the Chilean government implemented in 2008, a high-stakes accountability program was seen as a way to make schools responsible for student performance on standardized tests. It assumed that more pressure would create better results and therefore improve the overall quality of education. In this sense, Chile is a unique case because it has implemented the full platform of market-oriented reform, which includes competition, free choice, and the high-stakes accountability system. The question that arises is the following: What are the macro-level outcomes of these policies? One of the major problems in studying the impact of market-oriented policy in education is the lack of data. In this sense, Chile offers major advantages in studying the

impact of neoliberal policies because the government has built an extensive database tracking student performance (Carnoy & McEwan, 2000).

This dissertation studies the impact of competition and the accountability system in Chile. I use three objectives to do so. The first objective studies the intended and unintended consequences of the introduction of competition on student achievement. The second objective is focused on the intended consequences of the implementation of the high-stakes program, and in particular whether teachers are responding to the new accountability system by increasing their use of academic strategies; for example, how teachers are working harder, more, and finding better teaching methods to improve student leaning to improve the school's ranking. The third objective explores the unintended consequences of the new accountability program, specifically, it studies how teachers are increasing their use of non-academic strategies, such as the exclusion of low-achievement students, as a way to improve school performance on standardize tests.

I used qualitative and quantitative methods to explore these objectives. In the next section, I present the major findings of this study, summarized by chapters, and then the policy implications of this study are laid out. Finally, the limitations of the study and future research possibilities are described.

MAJOR FINDINGS

Chapter 3 analyzes the impact of competition on student achievement and how private-voucher schools can have a different impact depending where they are located. One of the main problems in estimating the causal relationship between competition and student achievement is endogeneity between these variables. For example, if a municipality provides a bad quality of education, private-voucher schools could have incentives to open and compete with public schools. Therefore, by using a traditional

regression model, researchers could erroneously conclude that an increase in competition has a negative effect on students achievement. By contrast, if a municipality has high quality education and private-voucher schools open to compete with public schools, researchers could find that competition has a positive effect on student achievement. In order to address this issue, the most recent studies in Chile have used instrumental variables on two-stage least square models (Auguste & Valenzuela, 2004; Gallego, 2002; Hsieh & Urquiola, 2003). However, none of these studies have been published in peer reviewed journals and the use of the instrumental variable has been criticized by Chilean scholars (Bellei, 2007). Instead, by using a longitudinal dataset, I am able to run regression models with fixed effects at the municipal level, which means that each municipality is compared with itself through time and the endogeneity problem is addressed by using this type of model.

The regression model shows that competition has a negative effect on student performance on the standardized test at a national level, but in the metropolitan region – where Santiago city is located – competition has a positive impact on student performance. Using research questions, I will present in more detail the main findings of Chapter 3.

How does competition affect student achievement? At the national level, the relationship between competition and student achievement from 2002 to 2013 is negative. For example, for a one percent increase in competition, student achievement decreases by 0.06 points on the reading test and 0.08 points on the math test. Both coefficients are statistically significant, but the effect is small compared with the standard deviation of the SIMCE tests, which is 50 points. However, this result is opposite of that which advocates of market-oriented policy have argued, which from a neoliberal perspective have suggested that the whole system would benefit from increased competition in education,

as schools would have incentives to improve their performance in order to attract students and, therefore, the quality of education would dramatically improve if parents were allowed to choose freely between those schools (Hsieh & Urquiola, 2003).

Moreover, one of the main assumptions of market-oriented reform in education is that if parents could exercise their free choice, they would choose the best schools for their children and, therefore, that schools would be forced to improve their quality of education in order to enroll students. However, using the information from the questionnaires from parents with students in 4th grade, I found that parents identified the most relevant variable in choosing a school for their children to be "school proximity" rather than "quality of education" or "results on the SIMCE test." Furthermore, "school proximity" becomes more relevant as the socioeconomic status of families decrease. Moreover, private-voucher schools have a mechanism by which they select their student body according to the socioeconomic status of families, and also based on students' cognitive abilities. In fact, even though it a 2008 prohibited schools that participate in the high-stakes accountability from selecting their student body, though information on obtained by questionnaires taken by parents with children in 4th grade in 2012 that 34.4 percent of the parents whose children attend SEP schools indicated that their children went through a selection process before they were admitted to their current school.

One of the most observable outcomes of the introduction of free choice and competition in Chile has been socioeconomic segregation of the system: poor students attend public schools, while the middle class send their children to privative voucher schools and the wealthiest families enroll their children in private paid schools. As critical theory has pointed out, if the introduction of competition and free choice could have any positive impact in the quality of education, this effect will be "cancelled out" by an increase in the socioeconomic segregation of the educational system.

The second research question in this chapter asks: Does competition have similar effects on those students who are educated in the capital city as those educated outside of it? In order to answer this question, I explore the diverse impact of competition on two different regions that have experienced a similar degree of competition, but that display dissimilar socioeconomic and demographic characteristics. I compare the metropolitan region, where the capital city is located and the area that is the wealthiest in Chile, versus the Araucanía region, which is the poorest region with the highest percentage of the population being indigenous and rural.

In relation to the metropolitan area of Santiago, competition has had a positive impact on student achievement. For example, when the percentage of students enrolled in private-voucher schools increases by one percent in 4^{th} grade, student scores on reading increase by 0.1 points and by 0.15 points on the math test, with both coefficients being statistically significant, in contrast, competition has a negative effect in the Araucanía region, where if the percentage of students in 4^{th} grade that enroll in private-voucher schools increases by one percent, reading scores decrease by 0.23 points and by 0.27 points on the math test. Both coefficients are statistically significant.

I argue that the negative impact of competition in the Araucanía region is due to the negative externalities generated by the financial system for rural schools, which creates incentives for the private sector to open small private-voucher schools in rural areas. Their main goal in doing so is to capture additional funding offered by the Chilean government to small rural schools. As a result, the number of private-voucher schools, and competition, has increased in rural areas. The quality of education, however, has not.

Finally, the findings of this chapter provide evidence that in the urban districts, where families are economically more advantaged than the rest of the country, competition might have a positive effect on student achievement. In contrast, in those

regions that are more disadvantaged, such as the Araucanía region, competition and free choice do not improve the overall quality of education. In summary, this chapter offers empirical evidence that runs contrary to assumptions made by advocates of market-oriented reform in education, those who hypothesized that the breakdown of the government's educational monopoly and the introduction of competition and free choice in the system would improve the overall quality of the system in the long run. Instead, the introduction of the private-voucher schools only reproduces the differences already existing in the country.

Chapter 4 analyzes the use of academic strategies after the implementation of the high-stakes accountability program. Academic strategies are the expected responses of teachers under an accountability program with the goal of improving students learning and, as a consequence, student performance on standardized tests. The chapter follows the typology developed by Daniel Koretz regarding teacher responses under the high-stakes accountability system. Academic strategies established by Koretz include (a) teachers work more effectively: for example, they find better methods of teaching, (b) teachers teach more: teachers spend more time overall in school, and (c) teachers work harder: for example, they give more homework or harder assignments to their students (Koretz, 2008). Academic strategies are relevant and expected to be implemented by teachers and schools because they produce higher levels of achievement on substantive areas of the test and an unambiguously meaningful gain in scores (Koretz & McCaffrey, 2001). In the current literature in Chile, it is not possible to quantify how extensive the use of academic strategies is under the high-stakes program implemented in 2008. In this sense, the findings of this chapter are relevant.

I will present the findings of this chapter following the next research questions: after the implementation of the high-stakes accountability in Chile in 2008, have teachers

increased the use of academic strategies as a way to improve students' performance on standardized tests? The first type of academic strategy that is expected is that teachers will work more effectively. The analysis of teacher questionnaires show that teachers working in schools that participate in the high-stakes program (SEP schools) versus those that are not participating (non-SEP schools) show a very similar pattern in terms of feedback that teachers are giving to their students and the type of evaluations that teachers are using. Furthermore, with relation to teaching methods, teacher questionnaires show that teachers who work in SEP schools have decreased their use of methods focused on individual and group work with their students and that SEP schools are not using clear or special teaching methods. I argue that this is because there is no clear guideline in SEP schools about which teaching methods should be used by teachers to improve student learning. Schools are more focused on improving their performance in SIMCE tests, rather than on improving students' cognitive abilities.

A second type of expected academic strategy is that teachers from SEP schools would simply teach more, spending more overall time with their students. I tested this assumption, using data about the average time that teachers spent preparing their classes before and after the implementation of SEP program. Teacher questionnaires show that from 2006 to 2010, teachers from SEP schools increased their preparation time by 1.79 hours, which follows the expected direction, but non-SEP schools also showed an average increase of 1.87 hours after the implementation the accountability system. This result presents empirical evidence that the expected response of "teaching more" under the pressures imposed by high-stakes accountability program is on its own responsible for increasing the time that teachers spend with their students or preparing their classes.

A third type of academic strategy is that teachers would work harder, giving more homework or harder assignments to their students under the high-stakes accountability

system. The results show that from 2011 to 2013, the percentage of teachers that reported to "always" give homework to their students was the same (2.2percent), but responses from teachers at non-SEP schools increased from 3.7 to 5.1 percent in the same period. Using data gathered before the implementation of the high-stakes accountability system, the percent of teachers from SEP schools that reported to "never" give homework to their students increased 10 percent from 2002 to 2013, while non-SEP schools showed an upward trend of 5.1 percent in the same period. In brief, these findings go against the actions expected of teachers under a high-stakes accountability program, as SEP schools have not seen an increase in the percentage of teachers that report to "always" give homework, while those who reported to "never" give homework to their students increased more for SEP school than for non-SEP schools in the period between 2002 and 2013.

The second research question of this chapter asks what the difficulties and constraints are that the Chilean educational system imposes on teachers and schools administrators in order to increase their use of academic strategies under the high-stakes accountability system. Based on content analysis of the semi-structured interviews, I argue that one of the main limitations that SEP schools face in improving their use of academic strategies is the leadership of principals, the quality of teachers, and the technical support that SEP schools receive. In relation to teacher training, for example, around 13 percent of teachers from SEP schools have been educated under special programs that represent the lowest levels of quality in the system (Ruffinelli & Rojas, 2007). In addition, the frequent rotation of teacher –as reported in the interviews – would not allow schools to develop educational projects in the long run while a lack a leadership among principals meant that resources were not used to promote educational projects at their schools. These factors, combined with the need to produce results, push teachers and

school administrators to seek strategies that produce results in the short run and that can be easily implemented; academic strategies do not offer these same characteristics.

In sum, Chapter 4 shows that teachers from SEP schools are not responding in the way that advocates of market-oriented policies have argued that teachers would. This means that there is empirical evidence that teachers from SEP schools are not increasing their use of academic strategies to improve student learning. Instead, the teacher questionnaires and the interviews offer evidence that SEP schools are shifting resources, including time, to emphasize the subjects and type of questions present on the test (reallocation), matching the curriculum more closely to the material covered on the test (alignment), or training students to take the standardized tests (coaching). SEP schools are under a lot of pressure to produce better results on standardized tests and the use of reallocation, alignment, and coaching seems to be more attractive than academic strategies while bringing them better results in the short term.

Chapter 5 explores the unintended consequences of the implementation of the high-stakes accountability system in 2008. In particular, I evaluate if teachers who work in SEP schools react to pressure intended to produce better results on standardized tests by using non-academics strategies – for example, excluding low-achieving students from the test-taking pool. This chapter addresses two main research questions, the first one asks: are schools, after participating in the high-stakes accountability program, more likely to exclude low-performing students from the test-taking pool in order to improve their overall scores on standardized tests? To answer this question I ran a logistic regression model with fixed effects at the school level, with an interaction term to measure the relationship between schools that participate in the high-stakes program (SEP schools) and low-performing students. In 4th grade, the results show that after the implementation of the SEP program, low-achievement students are 27.3 percent more

likely to be excluded from the reading test and 35.7 percent more likely to be excluded from the math test when their schools participate in the high-stakes program. Both coefficients are statistically significant. In the case of 8th grade, the logistic regression model shows that low-achievement students are 53.3 percent more likely to be excluded from the math test after their school begins to participate in the SEP program. The coefficient for the reading test is not statistically significant.

The second research question asks, how would the schools' rankings based on the standardized tests change students excluded from the test-taking pool were able to take the standardized test? To address this question, I use multiple imputation models in order to estimate the missing values in math and reading and recalculate the schools' rankings based on the SIMCE test. The results shows that in 2012, at the 2nd grade level, the schools' mean decreases on average by 1.42 points on the reading test when imputed values are included in the calculation. Furthermore, in 4th grade in 2012, the schools' means declines on average by 1.22 points on reading and 1.36 points on math when the imputed values of those students who did not take the standardized tests are included. Finally, after I use the imputed data to recalculate the schools' ranking in 4th and 8th grades, the percentage of schools classified as "insufficient" increased while schools ranked as "intermediate" and the percentage of schools classified as "advanced" decreased.

In summary, this chapter shows that after the implementation of the high-stakes accountability program, or SEP program, the likelihood that low-performing students would be excluded from the test-taking pool increased in 4th and 8th grades. Moreover, when the missing values from those students who did not take the reading or math tests are imputed and used to recalculate schools' mean in reading and math tests, the overall performance of the school goes down, the number of schools classified as "insufficient"

increase, and the number of schools ranked as "advanced" decline. There are issues of generalization in how standardized tests results are reported that could skew the evaluation of certain groups of students. This is the case, for example, with students from low-income families or racial minority groups. Chile could be underestimating the needs of certain groups because policy-makers are using biased information in the decision making process.

POLICY IMPLICATIONS

This study has been able to quantify the negative impact that competition has on student achievement at the national level: the effect is small but statistically significant, while teacher responses under the high-stakes program include a combination of academic and non-academic strategies. Several policy implications can stem from these results and they can be summarized as needs in the follow areas: (a) the strengthening of public education (b) improved teacher and principal training, and (c) the use standardized tests.

(a) The strengthening of public education

The implementation of private-voucher schools introduces competition in the system, but the actors don't compete under the same conditions. On one hand, private-voucher schools have been able to select their student body using mechanisms such as a selection test, interviews with parents, or by increasing school tuition. As a consequence, private-voucher schools are in an advantaged position to compete against public schools, because they can choose the type of student they want. Instead, public schools must admit all students. They cannot select their students and public schools are not able to charge additional tuition to parents. As a result, public schools are mostly attended by students from low-income families, and they have the reputation of being the worst schools in the

system. This has had an important effect on enrollment in public schools; in fact, from 2000 to 2013, public schools have seen a decrease in the percentage of students enrolling for primary and secondary education, from 54.6 percent in 2000 to 37.5 percent in 2013.

In January, 2015 the government approved an important law that prohibited any school that receives public funding from selecting students under any criteria while mandating that private-voucher schools be non-profit organizations. This law moves in the right direction, attempting to create a more even playing field in which public schools would compete against private-voucher schools. But a similar restriction was implemented in 2006 after the "penguin revolution," when schools were banned from selected students in elementary school. And as of 2008, schools that voluntary participate in the high-stakes program (SEP program) were no longer allowed to select students. However, from the questionnaire answered by parents with 4th grade children in 2012, I found that 34.3 percent of parents whose children attended SEP schools reported that their children went through a selection process before they were admitted to their schools. In this sense, the capacity of current institutions to supervise new rules and apply sanctions when they are violated is relevant. Otherwise, public schools could continue to lose enrollment and eventually they would disappear from the system, despite regulations that are intended to protect them from unfair competition with private-voucher schools.

Moreover, public schools are administrated by municipalities, and there are important differences among them in terms of available economic and human resources. For example, the departments of education of municipalities located in the capital city or in big urban sectors have more money and better human resources than those municipalities located in poor or less urban communities. In this sense, public schools also face an additional constraint in competing with private-vouchers schools, as they don't receive sufficient technical support from their municipalities that would allow them

to improve student learning. Instead, most of the municipalities are focused on the financial administration of public schools rather than the technical support offered to them. Therefore, it is of central importance that public school administration o does not remain a function of municipalities; instead, the government needs to create a decentralized institution that can support and give technical assistance to public schools.

This study offers empirical evidence that competition has a limited impact on student standardized test scores, and in contrast to what advocates of the market-oriented expected, competition among schools has had a negative impact on student achievement and it has also increased socioeconomic segregation of the system. As a consequence, the government, rather that creating or modifying rules to increase the competition among schools, should be focused on strengthening public schools and creating conditions for them to cooperate, reproducing pedagogical methods that have been shown effective in improving student learning. As this study has shown, competition by itself is not a solution and in the case of Chile, it has created more problems than benefits.

(b) Teacher and Principal Training

The main objective of the implementation of the high-stakes accountability program stems from the idea that teachers and schools do not face strong sanctions or pressure to improve student learning. Therefore, the idea is that if the results of student performance on standardized tests are linked with rewards or sanctions, student learning would improve because teachers would have incentives to do so.

This study found that it is not clear that schools participating in the high-stakes accountability program have increased the use of academic strategies, such as teachers working more effectively, teaching more or working harder, as a means of improving student achievement. One of the main limitations that SEP schools face in implementing

academic strategies is the enormous pressure to produce good results on SIMCE tests and the quality of teacher training. Many teachers are not prepared to deal with disadvantaged students or those with a different cultural background. In this sense, the Chilean government must regulate the conditions under which teachers are trained by universities, and stipulate the main conditions teachers need to satisfy in order to hold work in the public and also private sector.

Finally, one of the critiques of the implementation of the SEP program is related to the lack of connection between school spending and a school's educational project. In this sense, principals do not have a clear notion of how they should use new resources received from the SEP program to improve student learning. Therefore, it is important to improve the way that principals are selected and the requirements for their job, because under the new program, schools will have more flexibility and resources to spend, but if principals do not know how to link the new resources to the strengthening of the school's educational project, the new resources will have a limited impact on student learning.

(c) The use of standardized tests

In the context of the high-stakes accountability program, test scores have not been used to help schools. Instead, standardized tests are being used to sanction or punish schools. Schools face pressure and the threat of possible sanctions related to their ability to produce good results on the SIMCE test, and one way to improve their performance is by manipulating the test-taking pool. For example, Mizala & Torche (2013) estimated that schools that participate in the high-stakes accountability program have increased on average 4 points on reading and 5 points on math tests in 4^{th} grade. However, this study estimates that in 2012 schools that participate in the high-stakes accountability program the exclusion of low-performing students explain at least 1 point of the increase in

reading and math test in 4th grade. Moreover, I have found schools use non-academic strategies such as coaching, alignment or reallocation in order to improve students' performance. In consequence, the high-stakes accountability system probably is improving students' scores in standardized test, but not necessarily students' learning, what is the main objective of this type of educational policy.

Under the high-stakes accountability system, low-performing students are "undesirable" because they reduce average school performance. Students' academic performance is largely explained by a family's characteristics and, therefore, students from rural, indigenous, or poor families are more likely to perform lower than their peers from the middle class or those who have parents with a higher level of education. However, these students need to stay in the educational system, because they are the most disadvantages and education can make a difference for them.

LIMITATIONS

There are some limitations of this study and they fall into two broad categories. First, the years that were analyzed correspond to the last decade of implementation of the private-voucher schools, and this period only covers one third of the total period since privatization reform was implemented in 1981. The data available from standardized tests during the 1980s and early 1990s is not comparable with data from the 2000s because in 1996 there was a change in the methodology used to construct the SIMCE tests. Moreover, student-level data has only been available since 1999. This is the data that I am using to conduct this study. The second limitation is related to omitted variables in the models. For example, the logistic and regression models included the most relevant information about the student's family that, according to the literature, influences student achievement: family income and parents' education. However, it would be important

have data about whether or not students come from an indigenous family, if the student attended preschool, and to have the schools' financial information, among other variables. Furthermore, information relevant to the study of the use of academic strategies following the implementation of the high-stakes accountability system is not available in the teacher 'questionnaire. For example, it would useful to know how much time teachers spend working with their students, what their teaching methods are, to have an evaluation used before and after the implementation of the accountability system, and information about teachers training – for example, if the teacher attended e-learning, a part-time or full-time program.

FUTURE RESEARCH

One of the most important aspects of this study is the potential for future research, particularly in the area of high-stakes accountability system. In 2011, the Chilean Congress established in Law #20,529 that the voluntary high-stakes program implemented in 2008 would be mandatory for all schools in 2015 or 2016. What lessons can be learned from the SEP program implemented in 2008? How can the implementation of the mandatory accountability program, based on the weaknesses and strengths identified under the SEP program, improve? A number of questions arise, as do opportunities for future research in this area. I have identified three possible areas of research, looking at (a) how teachers respond to the high-stakes accountability system (b) the use of the standardized test to improve student learning, and (c) public schools: finance and support technical

(a) How teachers respond to the high-stakes accountability system

One area that is relevant to more deeply understanding this is knowing under what conditions schools and teachers decide to use academic or non-academic strategies to

improve student performance. Using the data from the SIMCE test, it would be possible to identify schools that participate in the program those that have improved their test scores since 2008. Then, one could use qualitative methods to explore what type of strategies these schools are using to improve student performance on the SIMCE tests. This could be an important source of information for the implementation of the mandatory high-stakes program, as the government could identify the conditions that facilitate teachers and schools choosing academics strategies while avoiding or reducing those that promote the use of non-academic strategies.

(b) The use of standardized test to improve pedagogical methods

One of the assumptions of the implementation of an accountability system is that schools and teachers will use this information to improve their pedagogical methods. However, schools are not currently using this information as an input to improve their work and if standardized test become more focused on supporting schools rather than penalizing them, as the current government has been proposed, it is fundamental to understand how this information should be delivered by the Chilean government in order to make it useful to schools.

(c) Public schools: finance and support technical

An interesting area of research is related to public schools, particularly in learning how to strengthen public education. The law approved in January, 2015 was intended to regulate the system and restrict the role and the scope of actions taken by private-voucher schools. But if the Chilean government does not create conditions or change the rules under which public schools are operating, the number of public schools could decrease drastically. There is a consensus in Chile that it is necessary to strengthen public schools, but the question is how. How can the administration of municipalities be improved? How are public schools getting their funding, and how can the quality of education that public

schools are offering be improved? How can the technical support that public schools receive be improved? These questions, among others, could be addressed by researchers in Chile.

SUMMARY

This study has evaluated the macro-level impact of two central aspects of market-oriented education policy: competition and accountability. While neoliberal policies have been highly promoted in the arena of education, there is limited evidence about their impact on the quality and equity of competition and the high-stakes accountability system.

The Chilean case shows that competition has had a negative impact on quality of education at the national level, but that families and students from Santiago (in the metropolitan region) have benefited from competition among schools, even though the effect is almost irrelevant in terms of public policy. However, one aspect that is significant and visible at the national level is the increase of socioeconomic segregation of students in the educational system as a result of the addition of private-voucher schools to the system. Furthermore, this study has analyzed how teacher responses to the implementation of the high-stakes accountability system in 2008, as well as the results, show that in those schools that participate in the high-stakes accountability system, in contrary to the expected outcome, teachers are not increasing their use of academic strategies, such as spending more time with students, finding new learning methods, or giving students more homework or assignments as a way to improve student achievement on standardized tests. Instead, teachers and schools under pressure to improve their performance on the SIMCE tests are increasingly using non-academic strategies, such as

excluding low-performing students from the test-taking pool, as a way to improve their performance on standardized tests.

While high-stakes accountability systems in education were developed according to the notion that some groups, such as students from low-income families or minority groups, typically receive poor service at their schools. As follows, it is also assumed an educational system that makes schools and teachers accountable for student learning would be especially beneficial to low-income families or minority groups. This is because, without explicit sanctions or rewards, advocates of accountability systems argue that teachers and principals will be less focused on the central objective of teaching the curriculum and improving student achievement. However, under continuing pressure to improve, established by an accountability system, schools are forced to use a combination of strategies that do not necessarily improve student learning, but rather help schools to artificially improve their performance on standardized tests and their ranking.

References

Auguste, S., & Valenzuela, J. P. (2004). *Do Student benefit from School Competition? Evidence from Chile*. University of Michigan.

Bellei, C. (2007). Expansion of private schools and educational improvement in Chile . *Revista Pensamiento Educativo*, *40*(1), 1–21.

Carnoy, M., & McEwan, P. J. (2000). The Effectiveness and Efficiency of Private Schools in Chile ' s Voucher System. *Educational Evaluation and Policy Analysis*, *22*(3), 213–239.

Gallego, F. A. (2002). *Competencia y Resultados Educativos:Teoria y Evidencia Para Chile* (No. 150). Santiago, Chile.

Hsieh, C., & Urquiola, M. (2002). *When Schools Compete, How Do They Compete? An Assessment Of Chile's Nationwide School Voucher Program*. *NBER Working Paper Series* (Vol. 10008). doi:10.1016/j.jpubeco.2005.11.002

Koretz, D. (2008). *Measuring Up: What Educational Testing Really Tells Us*. Harvard University Press.

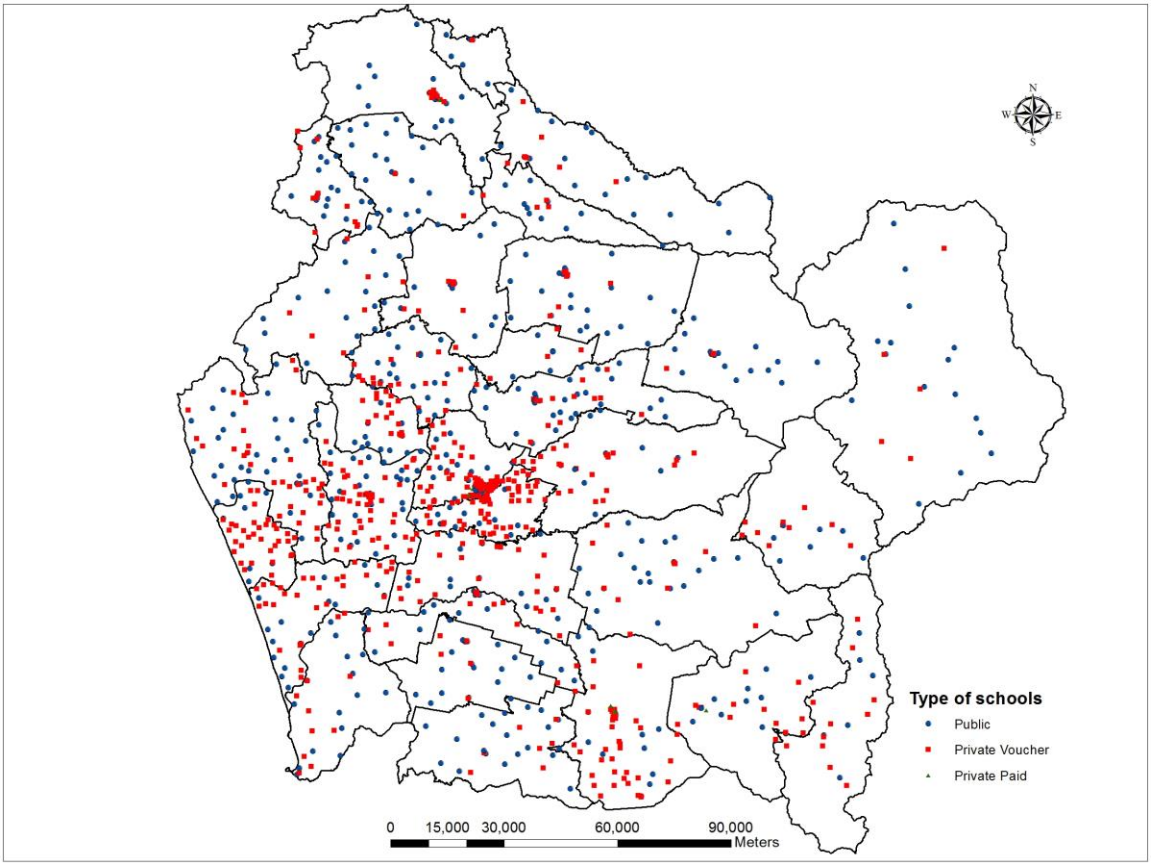
Koretz, D., & McCaffrey, D. (2001). *Toward a Framework for Validating Gains Under High-Stakes Conditions* (Vol. 1522). Los Angeles, CA.

Ruffinelli, A., & Rojas, M. T. (2007). *Sistematización de la oferta de Programas Especiales de pedagogía en Educación Básica de las instituciones de educación superior chilenas*. Santiago, Chile. Retrieved from http://cide.cl/asesoria/asesoria_desarrollo_detalle_actualizacion_sistema_ofertas_programas_pedagogia.php

Appendices

Appendix A. Competition

Figure A-1. Araucanía Region and Type of Schools in Primary Education 2011



Source: Author's calculations based on http://datos.gob.cl/datasets/ver/448

Table A-1. Fixed-Effects Model for 4th Grade Reading and Math Tests Only Including 12 Capital Cities of Each Region and Excluding Santiago City

Variable	Reading	Math
competition	-0.09***	-0.02
year		
2005	4.04***	0.25
2006	1.09***	-0.73*
2007	3.17***	-1.66***
2008	9.17***	-1.31***
2009	11.35***	5.24***
2010	19.42***	5.35***
2011	16.22***	11.97***
2012	16.79***	13.88***
2013	12.44***	7.35***
student's gender		
1= Male / 0=Female	-3.77***	9.95***
student's grades		
B	16.74***	18.27***
B+	38.43***	40.56***
A-	64.29***	68.31***
A	90.23***	97.35***
Father education		
High School	3.17***	3.12***
College	7.03***	6.87***
Mother education		
High School	2.63***	2.95***
College	5.84***	6.10***
Income of family (Monthly)		
US $401-1,000	1.41***	1.49***
US $1,001-2,000	1.42***	2.24***
US $2001-2,300	1.46***	2.63***
More than US $2,300	1.12**	4.06***
School type		
Private-voucher	5.48***	5.09***
Private-paid	3.96***	2.71***
Urban	-1.43***	0.56
School's SES		
Medium low	7.46***	9.52***
Medium	17.27***	20.85***
Medium high	24.25***	28.72***
High	29.41***	38.71***
_cons	187.78***	165.20***
N	335,412	336,582

Table A-2. Fixed-effects model that include dummy variable for those students who attend schools outside to their municipality

	National		Metropolitan Region		Araucania Region	
Variable	reading	math	reading	math	reading	math
competition	-0.07***	-0.09***	0.14***	0.15***	-0.14*	-0.24***
2008	base	base	base	base	base	base
2009	1.36***	6.05***	0.99***	5.24***	0.91	6.79***
2010	10.06***	5.65***	9.67***	4.49***	11.10***	8.88***
2011	6.23***	11.79***	5.01***	8.72***	9.53***	16.54***
2012	7.32***	14.75***	6.36***	11.67***	8.26***	18.01***
student's gender	-4.46***	11.29***	-3.94***	11.98***	-3.70***	11.95***
student's grades						
B	17.37***	17.81***	17.92***	18.44***	16.44***	17.08***
B+	39.07***	40.02***	40.05***	41.05***	38.77***	40.09***
A-	65.25***	68.34***	66.28***	69.41***	65.54***	69.48***
A	90.12***	98.18***	89.81***	97.99***	90.41***	100.54***
Study outside						
1= Outside	0.88***	0.55***	0.89***	0.92***	0.61	-0.59
Father education						
High School	2.54***	2.24***	2.69***	1.76***	2.86***	3.18***
College	5.30***	5.21***	4.63***	4.24***	6.40***	6.69***
Mother education						
High School	1.60***	2.51***	2.08***	2.39***	1.55***	2.89***
College	4.14***	5.21***	4.52***	4.75***	3.31***	6.16***
Income of family (Monthly)						
US $401-1,000	1.32***	1.66***	1.67***	2.14***	0.98*	0.56
US $1,001-2,000	2.09***	2.90***	2.55***	3.70***	0.89	2.55***
US $2001-2,300	2.69***	4.42***	3.58***	6.07***	0.34	1.93
More than US $2,300	2.95***	5.84***	3.99***	7.85***	-0.47	2.54*
School type						
Private-voucher	4.65***	3.91***	6.62***	6.05***	0.57	0.53
Private-paid	-0.09	-0.21	0.01	-1.07*	1.35	1.86
Urban	-0.92***	0.82***	-2.56***	-1.74***	3.72***	8.84***
School's SES						
Mediun low	5.72***	7.65***	8.16***	9.24***	8.05***	10.55***
Medium	15.99***	19.70***	20.69***	23.34***	15.87***	19.96***
Medium high	22.73***	29.21***	27.81***	33.36***	18.80***	27.37***
High	26.93***	35.61***	30.69***	38.44***	27.32***	40.42***
_cons	202.44***	174.55***	187.24***	161.78***	207.98***	174.69***
N	885,113	887,987	349,662	350,684	46,908	47,051

Appendix B. Interviews

B.1. Semi-structured Interview

B.1.1 Interviews with Teachers

*Implementación de la Subvención Preferencial Escolar

Narrativa que justifica el uso del SIMCE para clasificar a las escuelas para las escuelas

¿Por qué medir calidad?
¿Cómo se justifica el uso de pruebas estandarizadas para medir calidad?

Efectos negativos y positivos del uso del uso del SIMCE para clasificar a las escuelas

¿Cuáles son los efectos positivos del uso del SIMCE para clasificar las escuelas?
- ¿Se trabaja más tiempo preparando las clases?
- ¿Se utilizan nuevas estrategias pedagógicas para mejorar los aprendizajes de los alumnos?
- ¿Los profesores entregan más y más desafiantes tareas a los estudiantes?

¿Cuáles son los efectos negativos o no deseados del uso del SIMCE para clasificar las escuelas?
- ¿Las escuelas excluyen a los alumnos de bajo rendimiento?
- ¿Los estudiantes clasificados con necesidades educativas especiales transitorias rinden SIMCE?

*Relaciones laborales

¿Existe un ambiente de cooperación o competencia entre los docentes al interior de la escuela?
¿Cómo los directores ejercen presión sobre los docentes para obtener mejores resultados en el SIMCE?
¿Se siente apoyada por el director(a) de la escuela donde trabaja?
Si pudiese cambiarse de escuela a otra de similares características, ¿lo haría? ¿Por qué?

* Rol de Agencia de la Calidad de la Educación

¿Usted cree que la Agencia de la Calidad es una institución que ayudara a mejorar la calidad y equidad del sistema?
¿Cuáles serán los efectos no deseados de la implementación de un sistema de clasificación de las escuelas a nivel nacional?
¿Cuáles serán los efectos positivos de dicha implementación?

B.2. Sample of Transcript

Interviewed: Math teacher of Private-voucher school in Temuco City

Cuéntame fue tu experiencia como examinador del Simce

Las experiencias como examinador han sido eminentemente rurales, y estuve en 5 o 6 oportunidades como examinador en distintos colegios. Uno tiene que visitar los sectores rurales previamente, hacer un contacto donde se va a quedar, pues ahí no hay lugares donde uno pueda quedarse, están las escuelas no más. En general uno tiene que quedarse en la casa del director o algún profesor y cuando los lugares son muy lejanos te van a dejar un día antes y debe quedarse en la casa del director o profesores. En la mayoría de los casos ellos intentan conocer algo del proceso, pero el material viene sellado, a mí me han pedido en todos estos lugares si le pueden "echar una miradita" a la prueba, no se puede pues esta todo sellado"

En otro lugar un sostenedor, me dice que él tiene un conocido que era dueño de una escuela y que le ofreció una cantidad de dinero al examinador y a cambio de eso el examinador permitió que le ayudaran un poquito en la prueba. Insinuando indirectamente que podríamos llegar al mismo arreglo

En otra parte, la directora también directamente si yo pudiera ayudarles a los chicos, sabiendo que yo era profesor. Ella no pedía mucho, que uno le ayudara para que subieran 10 puntitos, 20 puntitos, directamente "*usted podría por favor ayudarle a los chiquillos*" así como de buena voluntad

En otro colegio, la experiencia que tuve fue que el primer día de aplicación, llegando a la sala, en un 4to básico, había un listado de alumnos en el pizarrón, eran 7 u 8. Y como había poco asistencia, yo les pregunte y que pasa chicos que no llegan los demás y estos que están anotados aquí quiénes son. Me dijeron "*esos son los porros, esos son los que no saben na tío y ellos vinieron en la mañana y el tío (profesor) se los llevo a Pucón en un furgón*". Los niños fueron a la escuela, pero el profesor los saco a pasear y eran como 7 u 8, era un porcentaje importante del curso

¿Y cuál ha sido su experiencia como profesor en donde ha tenido que ser evaluado por el Simce?

Mi experiencia como profesor, yo estoy siempre rotando pues en el sistema particular no te contratan por más de un año, entonces lo que hacen todos los colegios es focalizar

todos los esfuerzos en preparar a los estudiantes, con el tipo de formato del SIMCE, con el formato de ejercicio y como contestar la preguntas. Ocupando horas de clases que para ellos no son tan relevantes, que no sean matemáticas y lenguaje. Todo lo que es ciencias sociales, artes, música, tecnología, todos esos espacios de clases de esas asignaturas le hacen ensayos

¿Y qué pasa con el profesor de esa asignatura?

Él está ahí, pero hacen otra cosa o a veces también esta solo pero llevan ensayos. Entonces le dicen "*ya profe, aplique este ensayo de matemáticas*" y el profesor de música tiene que ir y aplicar el ensayo. Y eso es lo más recurrente en términos de la focalización del esfuerzo.

¿Y en tu caso como profesor de matemáticas tu sientes que hay presión por parte de los directivos o sostenedores hacia ti para que los niños obtengan buenos resultados en matemáticas?

Claro, no solo en matemáticas, se amplió el abanico de asignaturas. Ahora se rinde Simce en educación física, en ingles. Pero ellos están siempre más focalizado en lenguaje y matemáticas, y nos hacen presión en termino de resultados, pero no hay ningún cambio, siendo siempre lo mismo. Lo único que hacen ellos –directivos- es ocupar tiempo de otras asignaturas para ensayos, pero en tu asignatura no hay apoyo, no hay material extra con más tiempo, con más horas. Con lo mismo quieren que mejores los resultados

Te dicen esperemos que el próximo año el resultado sea mejor, pues usted tiene que preparar sus clases, tiene que mejorar. Porque si no hay mejora, usted sabe hay que buscar una nueva posibilidad de traer a otro tipo de profesional. Ósea, si usted no rinde usted no sirve

De hecho hace poco yo tuve un conflicto serio con la jefa de UTP, pues yo le dije que me sentía agredido, estaba un poco cansado con la Cantinela del Simce, porque todo los anos es la misma reunión, independiente en el colegio en que tu estés. Se reunen con los profesores de matemáticas y lenguaje y te dicen colegas en que estamos fallando, que estamos haciendo mal, que usted es el profesional, que está pasando no está haciendo su clase. Cambie su metodología, hágase una autocrítica... O sea somos los responsables, y yo le dije a la jefa de UTP, todos los años es la misma cantinela, que mejoren, que usted es el profesional, como si el resultado dependiera exclusivamente de la variable profesor

¿Y esto es generalizado o es una experiencia más bien personal?

Esto pasa en todos los colegios, uno se siente presionado, porque finalmente a ellos lo único que les interesa el índice que obtuvieron en el Simce

¿Y en las escuelas en que has trabajado cual es el nivel socioeconómico de los estudiantes?

Todos tienen un NSE deprivado (bajo), de hecho el año pasado el colegio tenía como el 90- 92 % de vulnerabilidad, y ahora por sobre el 80%. Pero es bien raro, porque los papas con los ingresos que declaran, porque declaran mayores ingresos, pero en realidad viven con mucho más que eso

¿Te han solicitado algún directivo de excluir a los alumnos de bajo rendimiento del Simce?

No, nunca así directamente no. Nunca me han pedido siendo yo profesor jefe de que yo trate que un estudiante no venga, pero siempre lo comentan igual informalmente "Ojala que estos cabros no vengan ese día" ese es un comentario del jefe de UTP o del director. Te sugieren que "ojala que estos cabros –así bien personalizado- no vinieran ese día porque nos van a echar a perder el Simce". Ese es un comentario bien común, pero nunca me han pedido así directamente como que hay que pedirle al papá o a la mamá que no los traiga ese día.

Este año el sostenedor para intentar cambiar un poquito la amenaza con un estímulo. Di ya, colegas si el próximo año subimos 10 puntos va a existir un bono, pero para el próximo ano, pero uno nunca sabe si el próximo año va a estar o no.

Cuéntame, ¿por qué es tan importante el SIMCE para las escuelas?

Por el ranking que se hace de establecimientos de la misma categoría, con los mismo ingresos, con el mismo tipo de dependencia. Se comparan entre ellos, y ahora también por los recursos SEP son extras tienen que ver con el rendimiento del Simce igual. Toda esa cantidad importante de recursos está en juego, entonces ellos te transmiten eso "colegas si nos va mal, no vamos a recibir este otro fondo" porque ellos se comprometen con los proyectos SEP a mejorar ese índice. Hay una presión económica

Entre los profesores y los directivos, como es la relación dado la presión que surge por mejorar en Simce?

Se genera una tensión permanente, en todos los colegios que he estado yo sucede exactamente lo mismo. O sea, de hecho los profesores no reunimos con los directivos después de la entrega del Simce y es siempre lo mismo, los profesores son los culpables del Simce. Porque tampoco asumen la responsabilidad como directivos, la dirección académica estuvo fallando

Muchos están presionados, genera angustia y estrés esta cosa del Simce. Es enfermiza esta cuestión dentro del colegio, pues produce división del equipo; tú tienes al cuerpo de profesores por aquí y los directivos acá. Uno atacando a los otros, y nosotros defendiéndonos y nosotros atacándonos. No hay espíritu de unidad de trabajar todos para el mismo lado

¿Y existe espacio de cooperación entre los profesores?

No hay espacio para eso, nosotros estamos abiertos a la posibilidad del trabajo en equipo y siempre buscando la instancia donde pudiese comentar con tu colega de matemática, yo creo que dentro de los profesores no hay problema con eso. Lo que sucede es que no hay tiempo o espacio para generar el departamento de matemáticas, te exigen pero no te dan espacio para coordinarte con los demás

Si invierten todo en reforzamiento, gente que venga de afuera, que trabaje con los estudiantes después de clases, que vengan los sábados, hasta para vacaciones de invierno se tomó un par de días para reforzamiento de Simce con profesores externos. Pero el profesor que trabaja en el aula, no hay ni un minuto más extra de tiempo, ni de pago

Y en relación al tipo de establecimientos, entre los municipales y los particulares subvencionados, ¿cuál de estos tendría mayor presión por mostrar avances en el Simce?

Yo no he trabajado en escuelas municipales, no tengo la experiencia. Pero creo que la presión siempre está en las particulares, ellas son el foco crítico ahora, pues se supone que debieran tener mejores resultados que los municipales. Y los municipales no se hacen mucho cargo, pues depende de toda una institucionalidad que hay en la comuna , el rendimiento tiene que ver con el conjunto de las escuelas municipales de la comunas, que con esa escuela en particular

¿Existe inestabilidad laboral para los profesores de escuelas particulares subvencionadas?

Absoluta, una cosa más que te agobia. Tu sabes que en Marzo tu tienes que comenzar a buscar trabajo de nuevo, entonces tampoco se genera un compromiso con el establecimiento donde uno vaya a decir voy a hacer un esfuerzo extra para este colegio le

vaya bien en el Simce porque el próximo año yo voy a seguir aquí, de hecho no hay ninguna seguridad, lo más seguro es que no sigas.

Yo he estado hasta dos años en un colegio, pero al segundo año si o si no te recontratan porque al segundo año por ley generas contrato indefinido. Y el contrato indefinido no le da la libertad al sostenedor de despedirte, porque te tienen que indemnizar y ya son tres años. Esta inestabilidad es transversal para todos los profesores que trabajan en el sistema particular subvencionado

Finalmente, ¿cuál crees tú que será el papel de la Agencia de la Calidad?

Yo creo que en ese escenario va a ser más complejo el tema, van a tener la amenaza no solo de subir sino también la amenaza del cierre. Entonces, van a intentar buscar los resultados independiente si los estudiantes son o no capaces de hacerlo, que el sistema logre hacerlo solamente por la vía formal y legal. Como hablar con el examinador, se puede agravar mucho más la situación porque será una presión extra.

Glossary

SIMCE	: Educational Quality Measurement System (Sistema Nacional de Evaluación de Calidad de la Educación)
SEP	: Preferential Voucher Program (Subvención Escolar Preferencial)
MINEDUC	: Ministry of Education (Ministerio de Educación)
CASEN	: Chilean National Household Survey (Encuesta de Caracterización Socioeconómica Nacional).
OLS	: Ordinary Least Square
MAR	: Missing At Random
MI	: Multiple Imputations

Bibliography

Ahlin, Å. (2003). Does School Competition Matter? Effects of a Large-Scale School Choice Reform on Student Performance. *Working Paper Series*. Retrieved from http://ideas.repec.org/p/hhs/uunewp/2003_002.html

Allison, P. (2002). *Missing Data* (Quatitativ.). Sage Publictaion.

Altonji, J. G., Huang, C., & Taber, C. R. (2005). *Estimating the Cream Skimming Effect of Private School Vouchers on Public School Students* (Vol. 1). Retrieved from http://www.ssc.wisc.edu/~ctaber/Papers/vouchpap.pdf

Andere, E. (2012). *The 2012 Education Reform in Mexico: From Intention to Action: The Expert Take*. Retrieved from https://mexicoinstitute.wordpress.com/2012/12/19/the-2012-education-reform-in-mexico-from-intention-to-action-the-expert-take/

Auguste, S., & Valenzuela, J. P. (2004). *Do Student benefit from School Competition? Evidence from Chile*. University of Michigan.

Avalos, B., Leyton, M., Soto, V., & Garcia-Huidobro, J. E. (2013, September 26). Carta abierta por un Nuevo Sistema de Evaluación Educacional. *El Mostrador*, p. 6. Santiago, Chile. Retrieved from www.elmostrador.cl/media/2013/Carta-abierta.pdf

Baker, E., Barton, P., Darling-Hammond, L., Haertel, E., Ladd, H., Linn, R., … Shepard, L. (2010). *Problems with the Use of Students Test Scores to Evaluate Teachers* (No. 278) (Vol. 03). Washington, DC.

Baker, G. (1992). Contracts Performance Measurement. *The Journal of Political Economy*, *100*(3), 598–614.

Baker, G., Jensen, M., & Murphy, K. (1988). Compensation and Incentives: Practice vs. Theory. *The Journal of Finance*, *XLIII*(3), 593–616.

Banco-Central. (2015). Producto Interno por Region. Retrieved June 2, 2015, from http://www.bcentral.cl/estadisticas-economicas/series-indicadores/index_aeg.htm

Belfield, C. R., & Levin, H. M. (2002a). The Effects of Competition Between Schools on Educational Outcomes: A Review for the United States. *Review of Educational Research*, *72*(2), 279–341. doi:10.3102/00346543072002279

Belfield, C. R., & Levin, H. M. (2002b). The Effects of Competition Between Schools on Educational Outcomes: A Review for the United States. *Review of Educational Research*, *72*(2), 279–341. doi:10.3102/00346543072002279

Bellei, C. (2007). Expansion of private schools and educational improvement in Chile . *Revista Pensamiento Educativo*, *40*(1), 1–21.

Bellei, C. (2012). *Equidad Educativa y Dinámicas de Mercado en Educación*. Santiago, Chile.

Benveniste, L., Carnoy, M., & Rothstein, R. (2003). *All Else Equal: Are Public and Private Schools Different?* NYC: RoutledgeFalmer.

Benware, C., & Deci, E. (1984). Quality of Learning With an Active Versus Passive Motivational Set. *American Educational Research Association*, *21*(4), 755–765.

Betebenner, D. W., Howe, K. R., & Foster, S. S. (2005). On School Choice and Test-Based Accountability. *Education Policy Analysis Archives*, *13*(41), 22.

Betts, J., & Loveless, T. (2005). *Getting choice right: ensuring equity and efficiency in education policy*. Washington, D.C.: Brookings Institution Press.

Booher-Jennings, J. (2005). Below the Bubble: "Educational Triage" and the Texas Accountability System. *American Educational Research Journal*, *42*(2), 231–268.

Borland, M. V, & Howsen, R. M. (1992). Student academic achievement and the degree of market concentration in education. *Economics of Education Review*, *11*(1), 31–39. Retrieved from http://www.sciencedirect.com/science/article/pii/027277579290019Y

Brian, J., & Levitt, S. (2003). Rotten Apples: An investigation of the Prevalence and Predictors of Teacher Cheating. *The Quaterly Journal of Economics*, (August), 843–877.

Brown, B. (1992). Why governments run schools. *Economics of Education Review*, *11*(4), 287–300. doi:10.1016/0272-7757(92)90038-5

Burgess, S., & Ratto, M. (2003). *The Role of Incentives in the Public Sector: Issues and Evidence* (No. 03). Bristol.

Cameron, J., & Pierce, W. D. (1994). Reinforcement , Reward , and Intrinsic Motivation: A Meta-Analysis. *American Educational Research Association*, *64*(3), 363–423.

Campbell, D. (1979). Assessing the impact of planned social change. *Evaluation and Program Planning*, *2*(1), 67–90.

Campbell, D. E., West, M. R., & Peterson, P. E. (2005). Participation in a national, means-tested school voucher program. *Journal of Policy Analysis and Management*, *24*(3), 523–541. Retrieved from http://dx.doi.org/10.1002/pam.20113

Carnoy, M. (2001). *School Vouchers: Examining the Evidence*. Washington, D.C.: Economic Policy Institute.

Carnoy, M., Loeb, S., & Smith, T. L. (2000). *Do Higher State Test Scores in Texas Make For Better High School Outcomes? The Civil Right s Project*.

Carnoy, M., & McEwan, P. (2003). *Choosing Choice: School Choice in International Perspective*. (G. Plank, David & Sykes, Ed.)*Comparative Education Review* (Vol. Vol. 42, N). New York: Teachers College Press.

Carnoy, M., & McEwan, P. J. (2000). The Effectiveness and Efficiency of Private Schools in Chile ' s Voucher System. *Educational Evaluation and Policy Analysis*, *22*(3), 213–239.

Carrasco, A., & San Martin, E. (2011). *Exploring New Approaches in School Choice Research: A Partial Examination of Educational Market-Driven Hypothesis From Value-Added Models*.

Carrasco, A., & San Martin, E. (2012). *Clasificacion de Escuelas en la Nueva Institucionalidad Educativa: Contribucion de Modelos de Valor Agregado para una Responsabilidad Justa* (No. 53). Santiago, Chile.

Casen. (2009). *Encuesta de Caracterización Socieconómica Nacional*. Retrieved from http://www.ministeriodesarrollosocial.gob.cl/casen2009/

Casen. (2011a). *Resultados Regionales Encuesta Casen 2011: Región de La Araucanía*. Retrieved from http://observatorio.ministeriodesarrollosocial.gob.cl/layout/doc/casen/IX_Araucania_oct13.pdf

Casen. (2011b). *Resultados Regionales Encuesta Casen 2011: Región de Metropolitana de Santiago*. Retrieved from http://observatorio.ministeriodesarrollosocial.gob.cl/layout/doc/casen/ResultadosRegionalesCasen2011_RegionMetropolitana_50e1bbf8db0b1.pdf

Casen. (2013). *CASEN 2013: Situación de la pobreza en Chile*. Retrieved from http://observatorio.ministeriodesarrollosocial.gob.cl/documentos/Casen2013_Situacion_Pobreza_Chile.pdf

Chapman, D. W., Wesley, C., & Jr, S. (2000). Can High Stakes National Testing Improve Instruction: Reexamining Conventional Wisdom. *International Journal of Education Development*, *20*, 457–474.

Contreras, D., Bustos, S., & Sepulveda, P. (2010). Emerging Evidence on Vouchers and Faith-Based Providers in Education. In F. Barrera-Osorio,

H. Patrinos, & Q. Wodon (Eds.), *Emerging Evidence on Vouchers and Faith-Based Providers in Education* (pp. 1349–1368). The World Bank.

Correa, J., Parro, D., Reyes, L., & Ugarte, G. (2012). *El impacto del voucher diferenciado en los resultados académicos: el caso chileno*.

Corvalan, J. (CIDE), Elacqua, G. (UDP), & Salazar, F. (UDP). (2009). *El sector particular subvencionado en Chile. Tipologización y perspectivas frente a las nuevas regulaciones*. Retrieved from http://biblioteca.uahurtado.cl/ujah/reduc/pdf/pdf/mfn341.pdf

Crouch, L. (2005). Political Economy, Incentives, and Teacher's Union: Case Studies in Chile and Peru. In E. Vegas (Ed.), *Incentives to Improve Teaching: Lessons from Latin America* (pp. 389–417). Washington, D.C.: The World Bank.

Cullen, B. J., & Reback, R. (2006). *Tinkering Toward Accolades: Schools Gaming Under a Performance Accountability System* (No. 12286). Retrieved from www.nber.or/papers/w12286

Deci, E., Nezlek, J., & Sheinman, L. (1981). Characteristics of the rewarder and intrinsic motivation of the rewardee. *Journal of Personality and Social Psychology*, *40*(1), 1–10. doi:10.1037//0022-3514.40.1.1

Deci, E., Ryan, R., & Koestner, R. (1999). A Meta-Analytic Review of Experiments Examining the Effects of Extrinsic Rewards on Intrinsic Motivation. *Psychological Bulletin*, *125*(6).

Deci, E., Ryan, R., & Williams, G. (1996). Need satisfaction and the self-regulation of learning. *Learning and Individual Differences*, *8*(3), 165–183. doi:10.1016/S1041-6080(96)90013-8

Deci, E., Schwartz, A., Sheinman, L., & Ryan, R. (1981). An instrument to assess adults' orientations toward control versus autonomy with children: Reflections on intrinsic motivation and perceived competence. *Journal of Educational Psychology*, *73*(5), 642–650. doi:10.1037//0022-0663.73.5.642

Dee, T. S. (1998). Competition and the quality of public schools. *Economics of Education Review*, *17*(4), 419–427. doi:10.1016/S0272-7757(97)00040-X

Dixit, A. (1997). Power of Incentives in Private versus Public Organizations. *The American Economic Review*, *87*(2), 378–382.

Dixit, A. (2002). Incentives and Organizations in the Public Sector An Interpretative. *Journal of Human Resources*, *37*(4), 696–727.

Dworkin, a. G. (2005). The No Child Left Behind Act: Accountability, High-Stakes Testing, and Roles for Sociologists. *Sociology of Education*, *78*(2), 170–174.

Elacqua, G. (2006). *Enrollment practices in response to vouchers: Evidence from Chile* (No. 125). New York City. Retrieved from http://www.ncspe.org

Elacqua, G., Santos, H., Martinez, M., & Urbina, D. (2011). *¿Estamos preparados para cerrar las malas escuelas en Chile? Impacto sobre equidad en el acceso a educación de calidad*. Santiago, Chile. Retrieved from http://www.comunidadescolar.cl/documentacion/FONIDE/Informe Final-Gregory Elacqua-UDP-511083.pdf

Eyzaguirre, B., & Fontaine, L. (1999). *¿Qué Mide Realmente el Simce?* (No. 75) (Vol. 75). Santiago, Chile. Retrieved from http://www.praxis.cl/sanisidoro/Tareas/02 Actividad de lectura/Qué mide realmente el SIMCE Eyzaguirre.pdf

Falabella, A. (2013). *Accountability Policy Effects Within School Markets A Study in Three Chilean Municipalities*. University of London. Retrieved from http://www.academia.edu/12249727/accountability_policy_effects_within_school_markets_a_study_in_three_chilean_municipalities

Falabella, A. (2014). The Performing School: The Effects of Market and Accountability Policies. *Education Policy Analysis Archives*, *22*(70), 1–29.

Falabella, A., & Opazo, C. (2014). *Sistema de Aseguramiento de la Calidad y procesos de mejoramiento: una mirada desde la gestión educativa*. Santiago, Chile. Retrieved from http://www.academia.edu/8568947/Sistema_de_Aseguramiento_de_la_Calidad_y_procesos_de_mejoramiento_una_mirada_desde_la_gesti?n_educativa

Fehr, E., & Falk, A. (2002). Psychological Foundations of Incentives. *European Economic Review*, *46*, 687–724.

Feltham, G. A., & Xie, J. (1994). Performance Measurement Congruity and Diversity in Congruity Principal / Agent Relations. *The Accounting Review*, *69*(3), 429–453.

Figlio, D. (2003). *Testing, Crimen and Punishment* (No. 11194). New York City. Retrieved from http://www.nber.org/papers/w11194

Figlio, D., & Getzler, L. (2002). *Accountability, Ability and Disability: Gaming The System* (No. 9307). Cambridge, MA.

Flores, C. (2008). *Residential Segregation and the Geography of Opportunities: A Spatial Analysis of Heterogeneity and Spillovers in Education*.

Fontaine, L., & Eyzaguirre, B. (2001). Una Estructura que Presione a las Escuelas a Hacerlo Bien *. In *¿Qué Hacer Ahora? Propuestas para el Desarrollo.* (pp. 66–111). Santiago, Chile: Centro de Estudios Publicos.

Friedman, M. (1962). *Capitalism and Freedom*. Chicago: University of Chicago Press.

Gallego, F. A. (2002). *Competencia y Resultados Educativos:Teoria y Evidencia Para Chile* (No. 150). Santiago, Chile.

Gallego, F. A. (2004). *School Choice , Incentives , and Academic Outcomes: Evidence for Chile* (No. 39). Retrieved from http://ideas.repec.org/p/ecm/latm04/39.html

Gallego, F., & Hernando, A. (2009). *School Choice in Chile: Looking at The Demand Side* (No. 356). Retrieved from http://www.economia.puc.cl/docs/dt_356.pdf

Gauri, V. (1998). *School Choice in Chile: Two Decades of Educational Reform*. Pittsburgh: University of Pittsburgh Press.

Gemello, J. M., & Osman, J. W. (1984). Estimating Private the Demand Enrollment for. *American Journal of Education*, *92*(3), 262–279. Retrieved from http://www.jstor.org/stable/1085005

Gobierno de Chile. Decreto con Fuerza de Ley 2 (1998). Retrieved from http://www.leychile.cl/Navegar?idNorma=127911

Gobierno de Chile. (2011). Plan Araucania 7. Retrieved December 10, 204AD, from http://araucania7.cl/el-plan/

Goldman, F. (2012, April 22). Camila Vallejo, the World's Most Glamorous Revolutionary. *New York Times*. Retrieved from http://www.nytimes.com/2012/04/08/magazine/camila-vallejo-the-worlds-most-glamorous-revolutionary.html?pagewanted=all&_r=0

González, P., Mizala, A., & Romaguera, P. (2004). *Vouchers, Inequalities and the Chilean Experience*. Santiago, Chile. Retrieved from http://www.ncspe.org/publications_files/Voucher_ChileOP94.pdf

Greene, J., Peterson, P., & Du, J. (1998). School Choice in Milwaukee. In *Learning from School Choice* (pp. 329–344). Brookings Institution Press.

Haertel, E., & Herman, J. (2005). *A Historical Perspective on Validity Arguments for Accountability Testing* (Vol. 1522). Loas Angeles, CA.

Hamilton, L. S., Stecher, B. M., & Klein, S. P. (2002). *Test-Based Accountability in Education*. Santa Monica, CA: RAND Ecation.

Haney, W. (2000). The Myth of the Texas Miracle in Education. *Education Analysis Policy Archives*, *8*(41). Retrieved from http://epaa.asu.edu/ojs/article/view/432

Hanushek, E. A. (1986). The Economics of Schooling: Production and Efficiency in Public Schools. *Journal of Economic Literature*, *24*(3), 1141–1177 CR – Copyright © 1986 American Econ. doi:10.2307/2725865

Hanushek, E. A., & Rivkin, S. G. (2003). The Economics of School Choice. In C. M. Hoxby (Ed.), (Vol. I).

Hanushek, E., & Raymond, M. (2005). Does school accountability lead to improved student performance? *Journal of Policy Analysis and Management*, *24*(2), 297–327.

Harackiewicz, J. M., Abrahams, S., & Wageman, R. (1987). Performance evaluation and intrinsic motivation: The effects of evaluative focus, rewards, and achievement orientation. *Journal of Personality and Social Psychology*, *53*(6), 1015–1023. doi:10.1037//0022-3514.53.6.1015

Haveman, R., & Wolfe, B. (1995). The Determinants of Children's Attainments: A Review of Methods and Findings. *Journal of Economic Literature*, *33*(4), 1829–1878 CR – Copyright © 1995 American Econ. doi:10.2307/2729315

Heckman, J. J., Heinrich, C., & Smith, J. (2002). *The Performance of Performance Standards* (No. 9002). Cambridge, MA.

Hirschman, A. O. (1970). *Exit, voice, and loyalty: responses to decline in firms, organizations, and states*. Cambridge, Mass: Harvard University Press, 1970.

Holmstrom, B., & Milgron, P. (1991). Multitask Principal-Agent Analyses : Incentive Contracts , Asset Ownership , and Job Design. *Journal of Law, Economics, & Organizations*, *7*(January 1991), 24–52.

Hout, M., & Stuart, E. (2011). *Incetives and Test-Based Accountability in Education*. Retrieved from http://standardizedtests.procon.org/sourcefiles/incentives-and-test-based-accountability-in-education.pdf

Hoxby, C. (1994). *Do Private Schools Provide Competition For Public Schools?* (No. 4978). Retrieved from http://www.nber.org/papers/w4978

Hoxby, C. (2000). Does Competition Among Public Schools Benefit Students and Taxpayers? *The American Economic Review*, *90*(5), 1209–1238. doi:10.1257/aer.97.5.2026

Hoxby, C. (2003). School choice and school competition: Evidence from the United States. *Swedish Economic Policy Review*, *10*, 9–65.

Hsieh, C., & Urquiola, M. (2002a). *When Schools Compete, How Do They Compete? An Assessment Of Chile's Nationwide School Voucher Program* (No. 43). New York City. Retrieved from http://www.nber.org/papers/w10008

Hsieh, C., & Urquiola, M. (2002b). *When Schools Compete, How Do They Compete? An Assessment Of Chile's Nationwide School Voucher Program*. *NBER Working Paper Series* (Vol. 10008). doi:10.1016/j.jpubeco.2005.11.002

Hursh, D. (2007). Assessing No Child Left Behind and the Rise of Neoliberal Education Policies. *American Educational Research Journal*, *44*(3), 493–518.

Izquierdo, S. (2013, October 21). ¿Eliminar el Simce? *La Tercera*, pp. 9–10. Santiago, Chile. Retrieved from www.latercera.cl

Jacob, B. (2005). Accountability, incentives and behavior: the impact of high-stakes testing in the Chicago Public Schools. *Journal of Public Economics*, *89*(5-6), 761–796.

Jacob, B. (2007). *Test-Based Accountability and Students Achievement: An Investigation of Differential Performance on NAEP and State Assessments* (No. 12817). Cambridge, MA.

Kamenetz, A. (2014). When Teachers, Not Students, Do The Cheating. Retrieved from http://www.npr.org/blogs/ed/2014/09/29/352494603/when-teachers-not-students-do-the-cheating

Kane, T. J., & Staiger, D. O. (2002). The Promise and Pitfalls of Using Imprecise School Accountability Measures. *Journal of Economic Perspectives*, *16*(4), 91–114.

Kim, S. (2004). Accountability and School Reform in the U . S . Public School System. *KEDI Journal of Educational Policy*, *1*(1), 2–35.

Koretz, D. (2002). Limitations in the Use of Achievement Tests as Measures of Educators ' Productivity. *The Journal of Human Resources*, *37*(4), 752–777.

Koretz, D. (2003). Using Multiple Measures to Address Perverse Incentives and Score Inflation. *Educational Measurement: Issues and Practice*, *22*(2), 18–26.

Koretz, D. (2008a). Further Steps Toward the Development of an Accountability-Oriented Science of Measurement. In K. Ryan & L. Shepard (Eds.), *The Future of Test-Based Educational Accountability* (pp. 71–91). Taylor and Francias, Hoboken.

Koretz, D. (2008b). *Measuring Up: What Educational Testing Really Tells Us*. Harvard University Press.

Koretz, D., & McCaffrey, D. (2001). *Toward a Framework for Validating Gains Under High-Stakes Conditions* (Vol. 1522). Los Angeles, CA.

Kremer, M., & Sarychev, A. (2000). *Why Do Governments Operate Schools?* Retrieved from http://post.economics.harvard.edu/faculty/kremer/papers.html.

Ladd, H. F. (2002). School Vouchers: A Critical View. *Journal of Economic Perspectives*, *16*(4), 3–24.

Laffont, J., & Martimort, D. (2001). *The Theory of Incentives: The Principal-Agent Model*. Princeton University Press.

Lankford, H., & Wyckoff, J. (1992). Primary and secondary school choice among public and religious alternatives. *Economics of Education Review*, *11*(4), 317–337. doi:10.1016/0272-7757(92)90040-A

Levin, H. (1998). Educational vouchers_ Effectiveness, choice, and costs.pdf. *Journal of Policy Analysis and Management*, *17*(3), 373–392.

Levin, H. (2002). A Comprehensive Framework for Evaluating Educational Vouchers. *Educational Evaluation and Policy Analysis*, *24*(3), 159–174.

Linn, R. L. (2007). *Considerations when selecting and combining measures. Paper presented at a workshop at the National Academies*. Washington, DC.

Madaus, G., & Clarke, M. (2001). The adverse impact of high stakes testing on minority students: evidence from 100 years of test data. In G. Orfield & M. Kornhaber (Eds.), *Raising standars or raising barriers? Inequality and high stakes testing in public education* (p. 51). The Century Foundation. Retrieved from http://files.eric.ed.gov/fulltext/ED450183.pdf

Manzi, J. (2014, March 17). ¿Ganamos algo eliminando el Simce? *La Tercera*, pp. 2–3. Santiago, Chile. Retrieved from www.latercera.cl

Manzi, J., Bogolasky, F., Grau, V., Guitiérrez, G., & Volante, P. (2014). *Análisis Sobre Valoraciones, Comprensión y Uso del SIMCE por parte de Directores Escolares de Establecimientos Subvencionados*. Santiago, Chile.

Mccaffrey, D., Lockwood, J. R., Koretz, D., & Hamilton, L. (2003). *Evaluating Models for Teacher Accountability*. Retrieved from http://www.cgp.upenn.edu/pdf/rand.pdf

McEwan, P., & Carnoy, M. (1999). *The Impact of Competition on Public School Quality: Longitudinal Evidence from Chile's voucher System* (No. Unpublished Manuscript).

McEwan, P. J. (2000a). *The Potential Impact of Large-Scale Voucher Programs*. *Review of Educational Research* (Vol. 70). Retrieved from http://rer.sagepub.com/cgi/doi/10.3102/00346543070002103

McEwan, P. J. (2000b). *The Potential Impact of Large-Scale Voucher Programs*. *Review of Educational Research* (Vol. 70).

Meckes, L. (2007). Evaluación y Estándares: Logros y desafíos para incrementar el impacto en calidad educativa. *Revista Pensamiento Educativo*, *40*(1), 351–371.

Mineduc. (2003). *Evaluación de Aprendizajes para una Educación de Calidad*. Santiago. Retrieved from http://www.agenciaeducacion.cl/wp-content/uploads/2013/02/Comision_Simce.pdf

Mineduc. (2005). Matricula por Estudiante. Retrieved January 2, 2015, from http://centroestudios.mineduc.cl/index.php?t=96&i=2&cc=2036&tm=2

Mineduc. (2007a). *Aplicacion de la Metodologia Para Establecer Puntajes de Corte en las Pruebas de Corte en las Pruebas SIMCE 4 Basico*. Santiago, Chile.

Mineduc. (2007b). *Guía de Subvenciones Educacionales*. Santiago. Retrieved from http://w7app.mineduc.cl/LiquidacionSubv/Principal/guia.pdf

Mineduc. (2008). *Manual de Uso de la Base de Datos SIMCE 2008 para 4 Basico*. Santiago, Chile.

Mineduc. (2009). *Primer Reporte de Resultados Nacionales SIMCE*. Santiago, Chile.

Mineduc. Ley 20.529: Sistema Nacional de Aseguramineto de la Calidad de la Educacion Parvularia, Basica y Media y su Fiscalizacion (2011).

Valparaiso. Retrieved from http://www.cned.cl/public/secciones/seccionsnac/normativa/LEY-20529_27-AGO-2011.pdf

Mineduc. (2012). *Impacto de la Ley SEP en SIMCE: Una Mirada a 4 años de su Implementacion*. Santiago, Chile. Retrieved from http://www.mineduc.cl/usuarios/bmineduc/doc/201208301544260.evidenciascem8.pdf

Mineduc. (2013a). *Base de Datos de Rendimiento Escolar 2002-2012*. Santiago, Chile.

Mineduc. (2013b). Matricula por Estudiante. Retrieved January 2, 2015, from http://centroestudios.mineduc.cl/index.php?t=96&i=2&cc=2036&tm=2

Mizala, A. (2007). La Subvencion Preferencial. In C. Peña & J. J. Brunner (Eds.), *La reforma al sistema escolar : aportes para el debate* (p. 129). Santiago, Chile: Universidad Diego Portales.

Mizala, A., & Romaguera, P. (2000). School Performance and Choice The Chilean Experience. *The Journal of Human Resources*, *35*(2), 392–417.

Mizala, A., & Romaguera, P. (2001). *Regulación, incentivos y remuneraciones de los profesores en Chile*. CEA. Retrieved from http://books.google.com/books/about/Regulacion_incentivos_y_remuneraciones.html?id=Jb6RYgEACAAJ&pgis=1

Mizala, A., & Torche, F. (2013). ¿Logra la subvención escolar preferencial igualar los resultados educativos? *Espacio Publico*, *09*, 1–36. Retrieved from http://www.espaciopublico.cl/publicaciones/detalle.tpl?id=19

Mizala, A., & Urquiola, M. (2007). *School Markets: The Impact of Information Approximating Schools' Effectiveness* (No. 13676). Cambridge, MA. Retrieved from http://www.nber.org/papers/w13676

Murnane, R. J., Page, L., & Vegas, E. (2010). La Subvención Escolar Preferencial (SEP) en Chile: Examinando las tendencias previas y su

primer año de implementación . Retrieved from www.dipres.gob.cl/594/articles-121390_doc_1pdf.pdf

Neal, D. (2011). How Vouchers Could Change the Market for Education. *The Journal of Economic Perpectives*, *16*(4), 25–44.

Neal, D., & Schanzenbach, D. (2010). Left behind by design: proficiency counts and test-based accountability. *The Review of Economics and Statistics*, *92*(May), 263–283.

Nechyba, T. J. (2000). Mobility, Targeting, and Private-School Vouchers. *The American Economic Review*, *90*(1), 130–146. Retrieved from http://www.jstor.org/stable/117284

Nechyba, T. J. (2003). Introducing School Choice into Multidistrict Public School Systems. In C. M. Hoxby (Ed.), *The Economic of School Choice* (pp. 145–192). University of Chicago Press.

Nichols, S. L., & Berliner, D. C. (2007). Collateral damage: how high-stakes testing corrupts America's schools. Cambridge, Mass.: Harvard Education Press. Retrieved from http://catalog.hathitrust.org/Record/005555685

Nicol, D., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning : A model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199–218. doi:10.1080/03075070600572090

Paredes, R., & Pinto, J. I. (2009). Is this the end of Public Education in Chile? *Estudios de Economia*, *36*, 47–66.

Perticara, M., Román, M., & Selman, J. (2013). *Rendimiento escolar y uso de los recursos de la ley SEP* (No. I-297). Retrieved from http://fen.uahurtado.cl/wp-content/uploads/2010/07/I-2971.pdf

Pyrczak, F. (1972). Objective evaluation of the quality of multiple-choice test items designed to measure comprehension of reading passages. *Reading Research Quarterly*, *8*(1), 62–71. doi:10.2307/746981

Raudenbush, S. W. (2004). *Schooling, Statistics and Poverty: Can We Measure School Improvement?* New Jersy. Retrieved from http://files.eric.ed.gov/fulltext/ED486417.pdf

Rawls, A. (2001). Eliminating Options Through Choice. *Emory Law Journal*, *50*(1), 363–396.

Robertson, S. (2000). *A Class Act: Changing Teachers' Work, Globalization and the State*. New York, NY 10001: Routledge.

Romaguera, P., & Gallegos, S. (2010). Financiando la Educación de Grupos Vulnerables: La Subvencion Escolar Preferencial. In O. Larrañaga & D. Contreras (Eds.), *Las Nuevas Politicas de Proteccion Social en Chile* (pp. 1–41). Santiago, Chile: PNUD.

Roman, M. (1999). *Usos Alternativos del Simce: Padres, Directivos y Docentes* (No. 5). Santiago, Chile.

Rothstein, R. (2008). *Holding Accountability to Account: How Scholarship and Experience in Other Fields Inform Exploration ofPerformance Incentives in Education* (No. 04). Retrieved from http://s4.epi.org/files/2014/holding-accountability-to-account.pdf

Ruffinelli, A., & Rojas, M. T. (2007). *Sistematización de la oferta de Programas Especiales de pedagogía en Educación Básica de las instituciones de educación superior chilenas*. Santiago, Chile. Retrieved from http://cide.cl/asesoria/asesoria_desarrollo_detalle_actualizacion_sistema_ofertas_programas_pedagogia.php

Salinas, D., & Fraser, P. (2012). Educational Opportunity and Contentious Politics : The 2011 Chilean Student Movement. *Berkely Review of Education*, *3*(1), 17–47.

Stecher, B. M., & Barron, S. (2001). Unintended Consequences of Test-Based Accountability When Testing in " Milepost " Grades. *Educational Assessment*, *7*(4), 259–281.

Taut, S., Cortés, F., Sebastian, C., & Preiss, D. (2009). Evaluating school and parent reports of the national student achievement testing system (SIMCE) in Chile: access, comprehension, and use. *Evaluation and Program Planning*, *32*(2), 129–37. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/19026445

Torche, A., & Sapelli, C. (2002). *Subsidios al Alumno o a la Escuela: Effects en la Eleccion de Colegios Publicos* (No. 202).

Torche, F. (2005). Privatization Reform and Inequality of Educational Opportunity : *Sociology of Education*, *78*(Mmi), 316–343.

Urquiola, M. S. (2000). Demand Matters: School District Concentration, Composition, and Educational Expenditure. *The World Bank*. doi:10.2139/ssrn.244099

Valenzuela, J. P., Bellei, C., & De Los Rios, D. (2008). *Evolución de la Segregación Socioeconómica de los Estudiantes Chilenos y su Relación con el Financiamiento Compartido*.

Vasquez-Heilig, J., & Darling-Hammond, L. (2008). Accountability Texas-Style: The Progress and Learning of Urban Minority Students in a High-Stakes Testing Context. *Educational Evaluation and Policy Analysis*, *30*(2), 75–110.

Villarroel, G. (2012). *Mejoramiento en Resultados Académicos de la Educación Básica en Chile: ¿Primeros efectos de la ley de Subvención Escolar Preferencial (SEP)?* Universidad de Chile.

Volante, L. (2007). *Evaluating Test-Based Accountability System: An International Perspective*. Stockholm, Sweden.

Waissbluth, M. (2013, July 21). Las Pervesiones del SIMCE. *La Tercera*, pp. 1–9. Santiago, Chile.

Wall, D. (2000). The impact of high-stakes testing on teaching and learning : can this be predicted or controlled? *System*, *28*, 499–509.

Weinstein, J., Fuenzalida, A., & Muñoz, G. (2009). *La subvencion preferencial: desde una dificil instalación hacia su institucionalizacion*. Santiago, Chile. Retrieved from http://ww2.educarchile.cl/UserFiles/P0038/File/publicaciones/La_subvencion_preferencial_desde_una_dificil_instalacion_hacia_su_institucionalizacion.pdf

West, B. (1997). Education Vouchers in Principal and Practice: A Survey. *The World Bank Research Observe*, *12*(1), 83–103.

Winters, M. a., Trivitt, J. R., & Greene, J. P. (2010). The impact of high-stakes testing on student proficiency in low-stakes subjects: Evidence from Florida's elementary science exam. *Economics of Education Review*, *29*(1), 138–146.